

## A stain on Italian reforms

**Italy's principal funding agency has missed an opportunity to enhance the prestige of its institutes. In appointing its first crop of new directors, it has conspicuously avoided some candidates of the highest calibre.**

A year ago, it seemed as if Italy's stodgy National Research Council (CNR) was taking its reform programme seriously. It had drawn up plans to merge its 330 institutes and centres to 100 or so larger units of 'critical mass'. And, with great fanfare, it announced in December its intentions of recruiting directors to the new institutes from the international community — a radical change from its traditional tendency to recruit from its own ranks. Moreover, advertisements referred to the need for 'continual, original and sound' research experience. The CNR claimed that the new research directors were key to ensuring that the reforms delivered in terms of high-quality science.

Disappointingly, the first round of appointments, for the largest of the new institutes, suggests business as usual. Of the first 22 appointments selected last month, all are Italian scientists who have served as CNR directors. And there are examples of clear injustices, where scientists with significant scientific and research-management experience have been passed over in favour of those of lesser merit. In one case, a leading scientist was passed over in favour of another with one-tenth of his publication output in the past decade. Moreover, the average citation rate of the rejected candidate was three times that of the winner. There is no mention of what factors led to the choice.

What went wrong with the plan for the CNR to break out of its mould? One key issue is that it was not given a budget for its reform — essential for making the positions attractive to top researchers from other countries. Not surprisingly, few highly qualified foreigners applied. But why there were no appointments of Italian scientists from

outside the CNR deserves explanation, as do those aberrant choices.

Another problem is the way in which the selection process took place. The procedures were out of line with international norms. CNR president Lucio Bianco is responsible for the final decision, but he is advised by (and heads) his Consiglio Direttivo, a committee of eight academics from all disciplines, from medicine to law. These committee members vote on candidates for each new institute directorship.

The Consiglio Direttivo did turn to experts for help. It appointed special commissions for each directorship, comprising three experts in the relevant field. But very few outsiders were brought in to ensure truly independent advice. The first 21 commissions whose names are publicly available included only four foreigners.

Moreover, the *consiglio* asked the commissions to provide unranked short-lists for its less technically expert members to choose from. And non-unanimous cases were combined to be voted on as a group rather than case by case. This system leaves room for suspicion of vote exchange — 'vote for my candidate and I'll owe you a favour' — the bad habit that the reforms should have forced the CNR to lose.

If the CNR appointments continue along the lines of these first ones, then it needs to come up publicly with a plausible explanation for why it cannot attract, or appoint, to key positions scientists from beyond its own ranks. Under attack from the new centre-right government of Silvio Berlusconi for inefficiency, the CNR needs more than ever to prove that it stands for research excellence and can strive to achieve it by seeing beyond its own highly politicized horizons. ■

## Visionary experimental designs

**A collaboration marrying epidemiology and genomics should provide a much-needed boost to analytical rigour.**

The widespread astonishment following the unconfirmed allegation by the UK government that researchers spent five years accidentally testing cattle instead of sheep brains for BSE is encouraging. It shows that politicians and the public alike appreciate a basic tenet of experimental design: be sure of the identity of what you are studying.

It may be shocking to some, therefore, that biologists knowingly transgress this rule daily. Genome databases are polluted with incorrect gene functions that are mistakenly assigned through researchers' uncritical faith in the results of BLAST algorithms. Powerful 'black-box' software packages mean researchers may plug data in, and get results out, with little thought for the rationale and caveats of the in-between. Many reported associations between diseases and DNA variants specific to particular regions of the genome have also recently emerged as being spurious or irreproducible (see *Nature Rev. Genet.* **2**, 91–99; 2001). Here the explanation often seems to lie in researchers' unfamiliarity with the rigour needed in the statistical and experimental design of such population experiments.

The epitome of the required rigour is perhaps exemplified by the field of epidemiology, steeped in statistics and experimental design. Classical epidemiology has brought enormous strides in health research. Ironically, epidemiology itself suffers from a major flaw: the

end points that it correlates — largely, crude clinical symptoms — are at best surrogates of the underlying biological basis of the disease.

Most of what we call 'diseases' are a kaleidoscope of conditions, with distinct origins, prognoses, risk factors, genetic susceptibilities, and responses to therapy. Until now, epidemiology has of necessity investigated a disease as if it were 'one' disease, whereas many variants of it may respond differently to the factors under study — a major confounding variable. Moreover, even in the most intensively studied diseases, identified risk factors account for only a fraction of the variation in morbidity and mortality. Much remains to be discovered.

Research would be substantially more effective if it could better identify patients with subtypes of a disease, and acquire a better understanding of the underlying biological correlates. That is the lofty goal of a new European project to marry high-throughput post-genomic technologies and epidemiology in a systems biology approach dubbed 'genomic epidemiology' (see page 139).

Funding for the project is uncertain, and technological obstacles abound, but it deserves support. The scientists behind it are showing vision by thinking outside their disciplinary and institutional boxes. In marrying epidemiology and high-tech post-genomics, they may not only rejuvenate epidemiology, but also set a new standard for experimental design in biology. ■





**Action stations**  
Panel calls for yet another space-station overhaul  
p136



**Playing catch-up**  
George W. Bush's science adviser tries to get up to speed  
p137



**Mercury malady**  
Japanese government pursued by victims of pollution disease  
p138



**Cut to the bone**  
Museums worry that new law could harm fossil collections  
p139

# Geneticists' work in disarray as DNA-chip producer pulls the plug

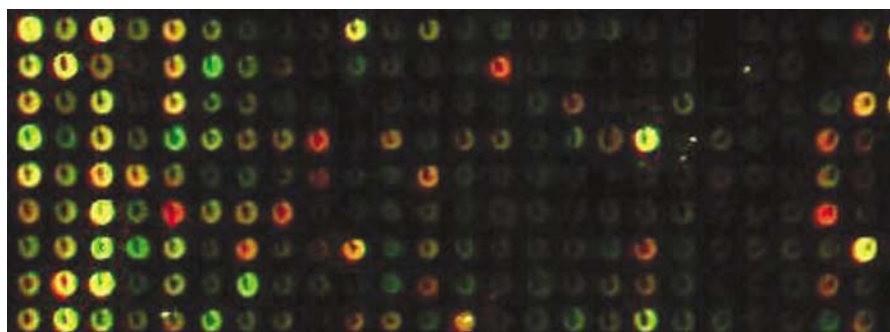
**Jonathan Knight, San Francisco**

A leading manufacturer of DNA chips has unexpectedly decided to get out of the business, leaving researchers scrambling to find substitutes for the company's products.

California-based Incyte Genomics announced on 24 October that it would stop making DNA chips, commonly known as microarrays, and end its gene-expression analysis service. It said the operation was no longer profitable.

It is thought that Incyte—one of the leaders in the emerging technology—is leaving the business in part because so many research institutions now make their own arrays. Incyte, which holds several key microarray patents, also authorized several companies to sell the chips and associated services.

The popularity of microarrays is currently soaring. They allow researchers to monitor the expression of many genes during process-



**Chips for the chop:** despite their popularity, microarrays are found to be unprofitable by Incyte.

es such as cancer development. An array consists of thousands of DNA spots, each corresponding to a gene. When RNA from tissue is processed and applied to the array, the spots corresponding to active genes light up.

Some Incyte users only learned of the company's decision from reporters. Among

them was Al George, director of genetic medicine at Vanderbilt University Medical School in Nashville, Tennessee, whose lab recently enlisted Incyte to develop a complementary-DNA library. He was surprised because Incyte had invested so much in creating collections of mouse and human genes, known as expressed sequence tags (ESTs), which can be deposited on glass slides to make microarrays. "They probably have one of the largest collections of ESTs on the planet," he says.

Julia Ljubimova, a cancer researcher at Cedars Sinai Medical Center in Los Angeles, says she was in the middle of a project using Incyte microarrays when the announcement came. "My project was destroyed," she says. Bruce Aronow, director of genome bioinformatics at the Children's Hospital Medical Center of Cincinnati, is also surprised by the news, and he notes that Incyte's microarrays were among the most reliable on the market.

But maintaining and replicating large collections of genes involves time-consuming quality control (see *Nature* **410**, 860–861; 2001), and the expense may have led to shrinking profits for Incyte.

This would have been compounded by the proliferation of microarray facilities on almost every major research campus over the past two years. These typically operate with university funds or grants from the National Institutes of Health, cutting costs for academic labs doing microarray research.

James Sikela, a geneticist at the University of Colorado Health Sciences Center in Denver, says his lab used mouse arrays from Incyte for a series of experiments three years

## New Zealand says yes to GM trials

**Peter Pockley, Sydney**

New Zealand will go ahead with field trials of genetically modified (GM) crops, but will delay their commercial release until 2003 at the earliest, the government has announced.

The decision, which follows an intense national debate, is regarded as a victory by supporters of GM technology.

Prime Minister Helen Clark said on 30 October that the government would accept most of the recommendations of the Royal Commission on Genetic Modification, which this year completed an extensive investigation of the technology (see *Nature* **412**, 569; 2001). She said that GM technology would have a "key role" in New Zealand's agriculture and in emerging industries.

Opponents of GM technology had campaigned vigorously, attracting as many as 20,000 people to demonstrations. The government also came under pressure from its own six Maori members of parliament and from the seven environmentalists who hold the balance of power.

The government says it will allow the rights and spiritual beliefs of the indigenous Maori, who comprise 15% of New Zealand's four-million population, to be represented in the approvals process for GM research.

Mita Rininui, chair of the Labour Maori Caucus, is pleased about the inclusion of the Maori in the process. But he warns that they will oppose the use of transgenic technology. "To interfere with another life-form is disrespectful and another form of cultural arrogance," he claims.

Research minister Pete Hodgson says the debate over GM crops has "changed New Zealand", explaining that "Pakeha [non-Maori] and Maori started to understand different world views and scientists learned they cannot exercise vetoes over others".

The government says that field trials of GM crops can begin once they have been approved under strict new rules. But there will be a two-year hold on the commercial release of genetically engineered products, with the exception of vaccines and drugs. ■

ago, but could not afford to use them routinely. Now, because of a new microarray facility on campus, funded in part by the National Institute on Alcohol Abuse and Alcoholism, he is getting back into microarray work. "Cost is the driving force," he says.

Incyte seems to have concluded that there is little money to be made from DNA chips. In a statement, Roy Whitfield, its chief executive officer, pointed to "increasing competition and margin erosion". As a result of the decision, some 400 workers at the company's facilities in Fremont, California, and St Louis, Missouri, will lose their jobs.

Steven Gullans, who heads the microarray centre at the Brigham and Women's Hospital in Boston, says: "Microarrays are becoming more of a commodity. The only way to compete is on price." Alternatives are available for researchers in the middle of projects with Incyte. Microarrays made using Incyte's gene collections will still be available through third-party providers such as Motorola, Agilent and PerkinElmer, the company says.

But this has led to uncertainty over compatibility, as different companies use different printing technologies. Agilent, for example, uses an inkjet printing method rather than direct DNA spotting. Several appeals for advice on alternatives have appeared on a microarray listserv at the University of California, San Francisco.

Incyte users will not be left in the lurch, says Gullans. "For the typical user in the street looking to get their sample analysed, they can find somewhere to do it," he says. ■

## Critical report leaves NASA's station strategy up in the air

William Triplett, Washington

NASA should completely reform its management of the International Space Station (ISS), continue with its current, curtailed plan for the project, and come back in two years time for permission to restore some of the project's previous scope.

That is the recommendation of the ISS Management and Cost Evaluation Task Force, a blue-ribbon panel established by the space agency and the White House Office of Management and Budget to assess NASA's largest and most problematic programme.

If the station's crew is ever to rise to six or seven — the curtailed plan allows for only three — NASA must first prove it has the "credibility" to manage the expansion properly, says the task force. But the smaller version of the station "will not achieve the unique research potential of the ISS", the task force says, because three crew will have practically no time to run experiments.

The wait-and-see approach is the best of three options open to NASA, says the task force, which was chaired by Thomas Young, former president of defence contractor Martin Marietta, now part of Lockheed Martin.

One of the others is to build only the scaled-down version, with no possibility of returning to the original design. But this would have "a significant adverse impact on



**Brought to book:** NASA will need to curtail its ambitions for the International Space Station.

science", says Young. The third option, of building to the original design, is simply "not credible" within the agency's likely future budget, the task force found. The ISS has already cost US\$25 billion, plus billions of dollars in launch costs.

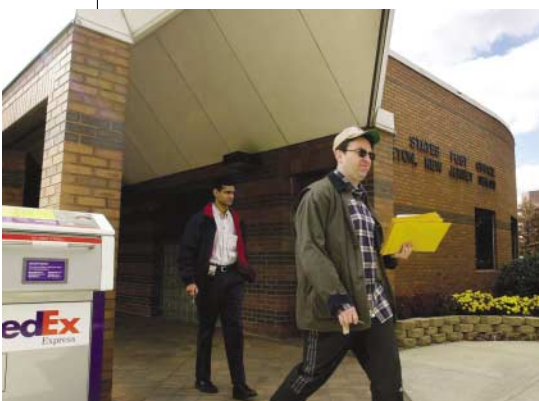
Even the first option may not be fiscally viable: it will cost \$500 million more over the next five years than the \$8.5 billion cap that Congress has placed on the project.

The task-force report contains a dozen recommendations on how to fix the project and control its costs, including the consolidation of all ISS-related staff and activities at NASA centres across the country under a single programme office. ■

## Universities address mail security as anthrax fears rise

Jonathan Knight, San Francisco

Worries about the mail being used for anthrax attacks have disrupted the normally benign security environments at university campuses across the United States.



**Sealed off:** staff leave the West Windsor post office after anthrax traces were discovered there.

Mail deliveries to Princeton University in New Jersey were suspended for two days from 27 October after anthrax spores were found in a mail-bin at the US post office in nearby West Windsor, which serves the campus. The building remains closed, but the post office began sorting mail in its parking lot.

An administrative building at Stanford University in California, meanwhile, was evacuated and closed twice, once on 22 October and again on 30 October, while suspect letters were tested. In each case, an employee reported finding a "suspicious substance", according to the university's own news service, accompanied by a threatening letter. Both samples tested safe and in each incident the building reopened the next day.

Security is becoming a top priority on most university campuses, as reports of suspicious envelopes and packages have become commonplace. For example, Ron Ozio, a spokesman for the University of

Pennsylvania in Philadelphia, estimates that 40 or 50 envelopes received there have already been tested for anthrax and other biological hazards, although none has come back positive.

Most university websites now include advice on how to deal with suspect mail, and Princeton is requiring mail handlers to take a special training course.

Some campuses have felt the impact of the anthrax attacks in the form of subpoenas or FBI visits. A spokesman for Louisiana State University in Baton Rouge confirmed that the FBI had asked anthrax researcher Martin Hugh-Jones to provide a list of all visitors to his laboratory since January 2000.

And officials at the Brookhaven National Laboratory in New York State received a subpoena asking them to testify in the US District Court in Miami regarding the lab's anthrax work, which has consisted of studies of a protein produced by the anthrax bacterium. ■



# Bush adviser vows to find science its voice

Matthew Davis, Washington

Ten months into a presidency that faces a raft of entirely unanticipated challenges at home and abroad, President George W. Bush finally has a science adviser.

John Marburger, the former director of the Brookhaven National Laboratory in New York state was confirmed in the post two weeks ago and has hit the ground running. He's already setting priorities for dealing with the deluge of terror-related research proposals submitted to the White House from various agencies. In addition, he is coordinating technical assessments of methods for improving anthrax detection and mail security, and is imposing a new organizational structure on the White House Office of Science and Technology Policy (OSTP), which he runs.

"Overall, we're getting up to speed after a long hiatus," says Marburger, who has been in Washington working as a consultant to Bush since late September.

But the mild-mannered physicist enters a White House playing-field where the only game in progress is the war on terrorism — and it's a contest that is already well under way. Marburger faces a stiff challenge, observers say, in asserting the influence of an office whose ranks have been depleted by neglect.

For example, the influence of a White House science adviser is normally measured by the degree of direct access the adviser has to the president. But on the most pressing scientific issue of the moment — bioterrorism — Marburger's advice will not go directly to Bush, but will instead be channelled through the newly created Office of Homeland Security and its chief, former Pennsylvania governor Tom Ridge.

Furthermore, Marburger is only one of several science advisers Ridge consults. And it has been suggested that Ridge — who will have broad powers in dealing with the anti-terror response — may appoint a chief science adviser of his own.

Marburger's influence on budgetary matters could also be limited, at least in the short term, as the 2003 budget proposal is already in its final phases of preparation.

He says he has been heavily involved in budget meetings on research spending, and has talked with the the Office of Management and Budget (OMB) about its plan to apply results-oriented investment criteria to science programmes. He pledges that the administration is "not seeking some crude sort of immediate pay-off" and that the OMB is "engaged in a serious effort at finding a better way of explaining to Congress and the public why basic, discovery-oriented science" is worth funding.

Marburger is in the midst of reorganizing the OSTP's staff structure. The associate



Tough target: John Marburger (right) needs to establish influence in the White House.

directorships for environment and national security are to go, leaving just two divisions: science and technology.

As for matters that occupied the White

House before 11 September, Marburger says he is looking at climate-change policy but has not spent much time thinking about human embryonic stem-cell research. ■

## Knowledge at stake in Australian poll

Peter Pockley, Sydney

Australia's opposition Labor Party has sought to make science, technology and education key issues in its campaign for the country's general election on 10 November.

Labor's call for more spending on research and at universities has gained support from unexpected quarters. In Melbourne last month, Australian-born



John Howard: "noodle nation".



Kim Beazley: tax incentives.

conservative media mogul Rupert Murdoch said: "We need urgent support for our centres of learning. It is no exaggeration to say we are threatened with global irrelevance."

But the warning was rejected by Prime Minister John Howard's governing right-wing Liberal Party, which looked set to lose the election earlier in the year. Howard is now polling well, after his action to prevent boatloads of asylum-seekers entering Australia and his offer of military help to the United States after the 11 September attacks.

The Labor Party's campaign centres on further funding of research and universities, the first time these issues have

attracted a lot of attention in an Australian election. Under a plan called 'Knowledge Nation' (see *Nature* 411, 619; 2001), leader Kim Beazley is pledging to double Australia's research spending, as a proportion of the economy, over the next 10 years.

Howard mocks the plan as "noodle nation" and says that his allocation last January of A\$2.9 billion (US\$1.5 billion) in new research funding over five years, was "the greatest ever single provision for science by any Australian government".

If elected, Labor promises to spend an extra A\$1 billion over five years on engaging more university lecturers. It would increase tax incentives for companies to use research at university and government laboratories, a move that would cost A\$88 million per year. The Commonwealth Scientific and Industrial Research Organisation, which has lost 1,000 staff since the Liberal government came to power in 1996, would receive an extra A\$160 million over five years.

But science leaders doubt that Labor's gradual provision of new funds will make its 10-year target attainable. Gavin Brown, chair of the Group of Eight universities (which dominate academic research), estimates that a further A\$13.6 billion will be needed to match the average research investment of the leading industrial countries by 2005.

Although Labor has declared its support for the current military action, analysts think that the issue of national security may favour the incumbent government. ■

## Drug price deal spells windfall for researchers

Xavier Bosch, Barcelona

Biomedical researchers in Spanish universities and hospitals are to receive a large cash injection from the pharmaceutical industry — in exchange for a promise from the government that it will not tighten price controls on drugs.

The unusual agreement will deliver US\$275 million over three years to the health ministry's research agency, the Instituto de Salud Carlos III (ISC). This figure almost matches the \$130 million a year that the Spanish government currently spends on biomedical research.

The announcement, made on 31 October, has been warmly welcomed by researchers. But although the money is supposed to come with no strings attached, some scientists have already expressed reservations that industry may influence how it is spent — and that the ISC may not distribute it efficiently.

The agreement is intended to ease the tension between the government and Farmaindustria, a body that represents 300 drug companies. Relations between the two have been strained since the health ministry lowered the caps that it imposes on drug prices by 6% in 1999.

In return for a promise from the government not to lower prices again, the industry will give the money to the ISC for "cancer, cardiovascular and genomics research". The industry also pledges to raise its annual budget for clinical trials conducted in public hospitals from \$80 million to \$135 million.

Of Spain's current investment in biomedical research, only about \$20 million comes from grants at the ISC. The agency has been supported by Farmaindustria before — last year it received \$30 million from the industry group to help build the Spanish National Cancer Centre (CNIO) and the Institute of Cardiovascular Diseases (IICV), both of which are under construction in Madrid (see *Nature* 406, 550; 2000).

The health ministry says that it has not decided exactly how to distribute the new funding. But some officials say that one-third of it will go to the CNIO, the IICV and genomics and proteomics research, with the rest going out as grants to research groups in public hospitals, and to support the training of researchers.

The government has rejected Farmaindustria's suggestion that a panel representing both parties should manage the money, and the industry is to set up its own panel to monitor how it is spent. ■

## Disputed diagnoses hamper claims of mercury poisoning



Deadly backwash: although now clean, Minamata Bay was a source of heavy mercury pollution.

David Cyranoski, Tokyo

A controversy over mercury poisoning that has dogged the Japanese government for nearly 50 years shows little sign of abating.

Minamata disease, a nervous disorder caused by methyl-mercury poisoning, was first identified in 1956. But at the sixth International Conference on Mercury as a Global Pollutant, held in Japan last month, experts were still arguing over how to diagnose the condition.

The waters around Minamata, a small town on the southern island of Kyushu, had been polluted since the 1950s with effluent containing mercury compounds from a chemical plant run by the company Chisso. Ever since the first cases of Minamata disease were identified, the issue of how victims should be compensated has caused the Japanese government problems.

A compensation scheme was finalized in 1995, offering victims certified as having the disease ¥2.6 million (US\$21,000) on the condition that they make no further claims and absolve the government of blame.

But some sufferers did not accept these terms and continued with their claim. And in April this year, the Osaka high court ordered the government and Chisso to pay US\$2.65 million to 51 plaintiffs who say they have the disease. The government is now appealing against this decision in the Supreme Court.

Minamata disease starts with numbness in the limbs, and progresses to slurred speech, hearing and vision disabilities, loss of coordination and, for many, death.

For years, one of the main tests for the disease gauged a patient's ability to register touch. A doctor would brush the tip of a patient's finger with a certain amount of pressure to see whether the patient could feel it.

But Shigeo Ekino, an anatomist at Kumamoto University, argues that such tests cannot distinguish between peripheral-nerve damage and damage to the central

nervous system, which the disease initially affects. Ekino's own research, presented at the meeting, uses a more sophisticated, two-point test in which two nearby spots on the finger are 'poked' with thread-like filaments.

Ekino, who gave evidence in the Osaka court case, believes that the government has underestimated the number of people suffering from the disease and that the true number could run into tens of thousands.

"The government's previous ways of judging whether someone was suffering from Minamata disease were unscientific and overlooked many victims," says Ekino, whose study compared residents of a fishing village on the same sea as Minamata with a control group in a similar village elsewhere in Japan.

But Akihiro Igata, director of the Aichi Comprehensive Health Science Center, who supervised the original testing of Minamata victims and was on the government's council for environmental health until January 2001, defends the old judgements. "We also did two-point discrimination tests, but we found no major difference," he says.

Ekino argues that Igata's tests were not sensitive enough. People who have suffered quite severe damage can distinguish poking with two filaments that are one or two centimetres apart, he says, but when the distance between them is reduced to about 3 millimetres, only those who have suffered damage to the central nervous system are unable to sense two distinct stimuli. The results of the debate could affect compensation claims by thousands of residents who are still applying for certification as Minamata victims.

Ekino also wants new surveys to establish the true extent of the disease, and claims that the government is not releasing data on mercury levels that it has already gathered. Environmental groups, meanwhile, are also pressing for the government to conduct health surveys and to give relief to Minamata victims in light of the new findings. ■



# Epidemiology set to get fast-track treatment

Declan Butler, Paris

A consortium of leading European research centres and pharmaceutical companies will this week announce a plan to transform epidemiology by combining it with the new techniques of high-throughput biology.

They plan to create a new field of study — 'genomic epidemiology' — by using screening technologies derived from the human genome project to chart the molecular, metabolic and disease profiles of thousands of clinical-trial subjects.

The goal is to correlate genetic markers with observations derived from a series of novel biological markers of disease and disease progression, using clinical trials on an epidemiological scale.

The consortium's 10 founding members include the University of Oxford, whose Nuffield Department of Clinical Medicine has access to huge amounts of trial data, and France's National Centre for Genotyping (CNG) near Paris, which specializes in high-throughput genotyping. The European Bioinformatics Institute near Cambridge will provide computing capabilities. Pharmaceutical companies Roche, Novo Nordisk and AstraZeneca are also participating. All data generated will be placed in public-domain databases.

The initial aim of the consortium will be to annotate tissue samples associated with subjects in trials concerning metabolic disorders such as adult-onset diabetes and cardiovascular disease. But the broader goal is to scale up technologies to epidemiological-sized samples, and to create generic tools and protocols that can then be applied to other diseases.

New funding for the multimillion dollar project is still being negotiated, but work can already proceed by redirecting and coordinating funds among the well-resourced participants, says CNG head Mark Lathrop.

A major limitation of current epidemiological studies is their reliance on clinical descriptions of 'disease', advocates of the new approach say. For example, in the 1950s hepatitis was seen as only one disease, albeit with a wide variety of symptoms and epidemiology. Now that the viral mechanism of the condition is understood, it is known that subtypes A, B, C, D, E and F exist and each have very different effects and prognosis.

"Many diseases are in fact a host of different diseases involving different mechanisms, hidden within the same terminology," says John Bell, head of Oxford's clinical medicine department. If certain factors, such as the environment, affect only one subtype, classical epidemiology might not detect that subtype.

The new approach aims to tackle this



problem by characterizing disease at a systems biology level, using genetics, proteomics, physiology and population biology, says Lathrop. It will collect unprecedented volumes of data on tens of thousands of

individuals and store this in shared databases. New tools will be developed to search these for patterns in gene, protein and metabolite data, which will be used to generate hypotheses for further research.

Genetics of disease will form an important aspect of the plan. But the consortium also aims to profile a subject's protein patterns using proteomic screens, and to take a quantitative analysis of every low-molecular-weight metabolite in various body fluids, using an emerging technology called 'metabolomics'.

"We think it is important to expand classical epidemiology and genetic epidemiology to take it to this high-throughput mode," says Esper Boel, vice-president of biotechnology research at Novo Nordisk. "We want to use post-genomic technologies to create a new clinical science, to turn functional genomics into real clinical chemistry." ■

## Museums choked by bone law

Rex Dalton, San Diego

A politically sensitive battle is brewing in California over the repatriation of Native American skeletal remains held in museums.

Scientists say that a state law enacted this month, which requires publicly funded institutions to make items available for claiming by Native American tribes, could deplete their fossil collections.

Passed with little opposition, the law was modelled on a federal statute. But the Californian Native American Graves Protection and Repatriation Act (NAGPRA) gives Native Americans more power in adjudications: tribal representatives will hold seven of the ten positions on a dispute-resolving commission. And unlike the federal statute, the state law allows non-federally recognized tribes to claim remains. California has more than 100 federally recognized tribes and at least 50 others.

The state commission's structure will lead to a "highly politicized" situation, says Phillip Walker, an anthropologist at the University of California, Santa Barbara. The federal body has a balance of scientists and representatives of Native Americans.

State Assemblyman Darryl Steinberg, who introduced the legislation, says it is right to favour Native Americans, because not enough material has been returned so far.

Tim White, curator of biological anthropology at Phoebe A. Hearst Museum of Anthropology at the University of California, Berkeley, is worried that questionable claims may succeed. The

museum has some 9,600 catalogued specimens, many from prehistoric humans.

Representatives of tribes who lobbied for the law declined to comment. ■



Pole-axed: staff at Harvard's Peabody Museum prepare to return a totem pole to Alaska.

ANGELA ROWLINGS/AP

## EPA finally agrees that arsenic levels in water should be cut

**Washington** After months of controversy, President George W. Bush's administration is reinstating a plan, inherited from its predecessor, to cut the amount of arsenic allowed in drinking water. The Environmental Protection Agency (EPA) says that the maximum legal limit for arsenic will be cut from 50 to 10 parts per billion from 2006.

In March, the EPA questioned the need for such a reduction, prompting criticism from experts in water quality (see *Nature* **410**, 503; 2001). The following month, the agency asked the National Academy of Sciences to review the issue.



**EPA's Christie Whitman: cuts are needed.**

"I said in April that we would obtain the necessary scientific and cost review to ensure a standard that fully protects the health of all Americans," EPA administrator Christie Whitman announced on 31 October.

"We did that and we are

reassured by all of the data that significant reductions are necessary."

## EU turns to satellites to monitor global security

**Munich** Environmental scientists and security experts should, from 2008, have access to a coordinated global monitoring system, courtesy of the European Union (EU). In Luxembourg on 30 October, research ministers from the EU's member states gave their backing to the Global Monitoring for Environment and Security (GMES) system, to be financed by the European Commission and the European Space Agency (ESA).

The GMES will combine information gathered by existing Earth observation satellites and ground-based sensors, some of which may be built specially for the system. Its scope, cost and technological framework will be agreed over the next two years by a joint industry-public consortium.

Some ministers expressed concern that the security focus, which has gained emphasis in the wake of the 11 September terrorist attacks on the United States, represented a move by the EU into military research. But research commissioner Philippe Busquin denied this. He argued that the system could help to detect movements of refugees, and so optimize the provision of aid. Environmental

applications include the management of natural disasters and monitoring of land use.

The GMES is part of an effort to coordinate EU space policy with the activities of ESA. It follows a joint decision earlier this year to create a global satellite navigation system called Galileo (see *Nature* **410**, 853; 2001).

♦ gmes.jrc.it

## Plan to back stem-cell research postponed

**Washington** Biologists hoping that the US Senate would try to roll back President Bush's restrictions on research using human embryonic stem (ES) cells will have to wait until at least February next year for the attempt. Senators last week postponed an attempt to allow federal funding to be used to create new human ES-cell lines.

A provision in the 2002 health-spending bill introduced by Arlen Specter (Republican, Pennsylvania) sought to overturn rules announced by Bush on 9 August, that government money will only be released for research on existing cell lines. But Specter withdrew the provision after Kansas Republican Sam Brownback proposed an amendment to ban funding for human ES-cell research altogether.

Specter and other supporters say they may re-introduce their proposal early next year.



## Sequencers reveal mystery of life on the ocean wave

**Washington** For the first time, researchers at sea have sequenced the DNA of marine organisms. By 1 November, the final day of a 17-day cruise, marine biologists from the University of Delaware had sequenced just under two million base pairs from the DNA of a selection of microbes and other organisms that live around intensely hot hydrothermal vents on the Pacific Ocean floor. That is about the same amount of DNA as is found in a small bacterial genome.

Technologies provided by Amersham Biosciences of Piscataway, New Jersey, were used to sequence the samples as soon as they were brought aboard the *Atlantis* from the

submersible *Alvin*. The project could give rise to new drugs, says Jim Yoder, director of the National Science Foundation's ocean-science division, which funded the research.

## Internet to link vast astronomy databases

**Washington** Astronomers are launching an ambitious plan to link dozens of databases across the Internet. The US National Virtual Observatory has been awarded a \$10-million grant from the National Science Foundation to present information gathered from both ground-based and orbiting telescopes in a unified form. A parallel initiative, called the Astrophysical Virtual Observatory, is being planned by the European Southern Observatory in Garching, Germany, and will receive 4 million euros (US\$3.6 million) in funding from the European Commission.

Advances in observation techniques and technology double the amount of data in astronomical databases each year. "This project will reach across the astronomical community," predicts Alex Szalay, an astronomer at Johns Hopkins University who is coordinating the US scheme with computer scientist Paul Messina of the California Institute of Technology.

♦ [us-vo.org](http://us-vo.org)

♦ [www.eso.org/projects/avol/index.html](http://www.eso.org/projects/avol/index.html)

## United States ready to talk about bioweapons control

**Geneva** The White House last week outlined the United States' negotiating position at the five-yearly review meeting of the Biological Weapons Convention, to be held in Geneva later this month. The statement has revived hopes of agreeing a protocol to enforce the ruling of the 1972 convention, following the US rejection of a draft protocol in July (see *Nature* 412, 365; 2001).

Of the seven points in the plan, two are taken from the protocol that the United States previously rejected. These points state that signatories should "enact strict national criminal legislation against prohibited biological weapons activities with strong extradition requirements", and call for "an effective United Nations procedure for investigating suspicious outbreaks or allegations of biological weapons use". Other points include "a code of ethical conduct for scientists", the scope of which is unclear.

But the United States remains opposed to mandatory inspection of plants where biological weapons might be manufactured.

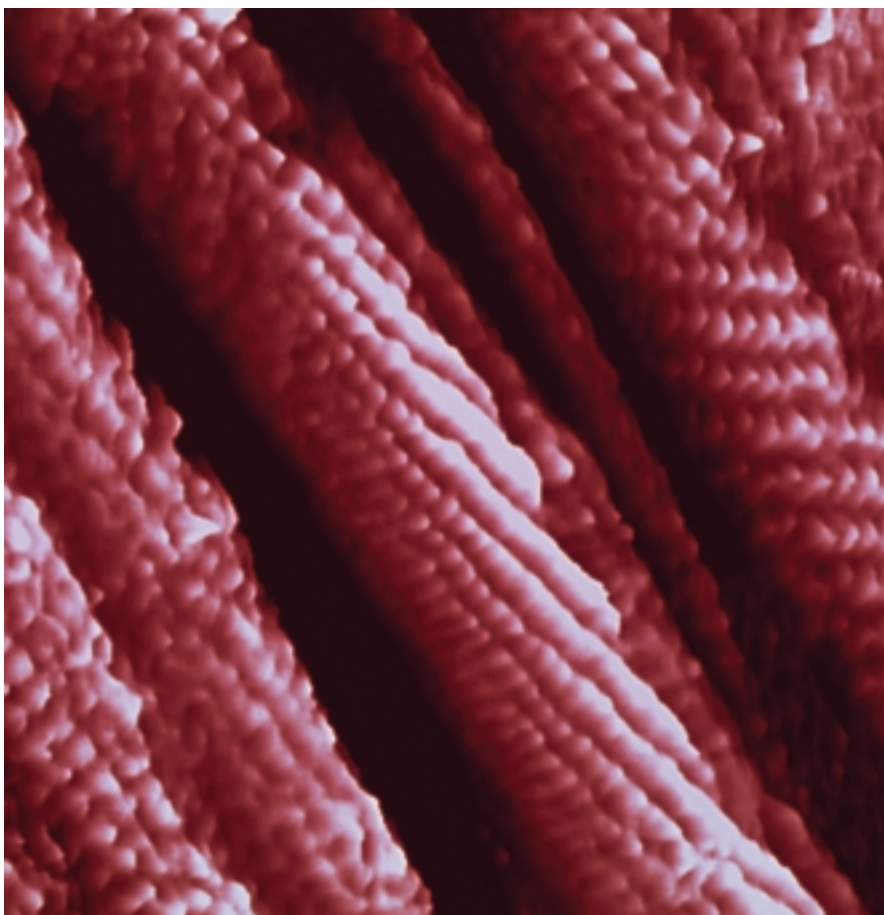
Jenni Rissanen of the Acronym Institute, a non-governmental organization that specializes in arms-control issues, says the proposals will be "welcomed as a basis for further work".



Nothing to hide: organisms such as the tube-living Pompeii worm are giving up their secrets.

# Roll up for the revolution

In the early 1990s, a chance finding in a Japanese laboratory introduced the world to carbon nanotubes. Today, interest in the tubes is still growing. Philip Ball reports on a decade of discovery.

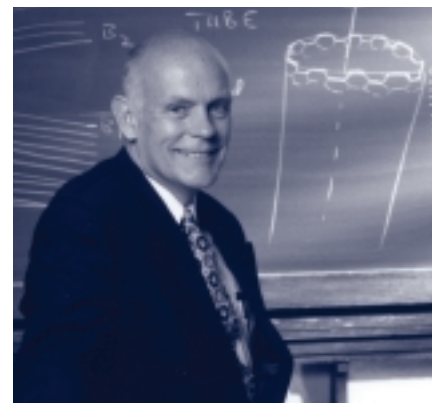
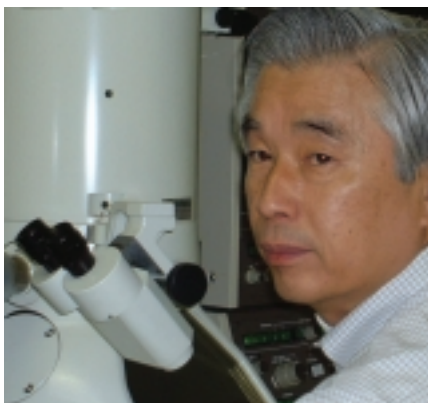


EYE OF SCIENCE/SPL

Ten years ago this month, Sumio Iijima, a microscopist at the NEC Corporation laboratories in Tsukuba, Japan, described how he found a needle in a haystack that was to revolutionize nanoscale science. Iijima's hollow needle, which had formed on the end of a graphite electrode, was just a few nanometres wide, several micrometres long, and made of pure carbon. It was a carbon nanotube<sup>1</sup>.

A decade on, nanotubes have accumulated an impressive list of real and potential applications. They have improved the precision of scanning probe microscopes<sup>2</sup>, and formed the basis for nanoscale tweezers that can pick up and move tiny particles<sup>3</sup>. In the past few months alone, logic gates<sup>4,5</sup>, the building blocks of computers, have been made from nanotubes, raising hopes that they could be used to create electronic circuits or even whole computers. "Nanotubes will be cheap, environmentally friendly, and do wonders for humankind," says Richard Smalley of Rice University in Houston, Texas.

Carbon nanotubes arrived at a time when physicists and chemists were exploring a new form of carbon — the cage-like fullerene molecules discovered in 1985 by Harry Kroto of the University of Sussex in England in collaboration with Smalley's team in Houston<sup>6</sup>. In graphite, carbon exists as flat sheets, with all the atoms arranged in a hexagonal pattern. But in fullerenes, there are also pentagons, so a sheet of atoms can form a hollow sphere.



In the pipeline: Richard Smalley (right) believes that carbon nanotubes (top), which were brought to the world's attention by Sumio Iijima (left), will be cheap and environmentally friendly materials.

In 1990, a team led by Wolfgang Krätschmer at the Max Planck Institute for Nuclear Physics in Heidelberg, Germany, and Donald Huffman of the University of Arizona in Tucson showed that fullerenes could be mass-produced by passing an electric current between two touching graphite electrodes surrounded by helium<sup>7</sup>. The heat produced vaporizes the graphite, and fullerenes form as the gaseous carbon cools.

Iijima had been studying carbon-fibre production, which uses similar methods, since the 1980s. Inspired by the mass production study, he tinkered with Krätschmer and Huffman's experimental conditions. Rather than bringing the two electrodes together,

Iijima kept them a short distance apart, so that sparks passed between them. To his surprise, he found fine carbon needles — nanotubes — growing on the negative electrode.

The nanotubes looked like sheets of graphite rolled up into tubes, with the ends closed off by caps of various shapes. But unlike fullerenes, the tubes were made up of a series of cylinders nestling one inside the other like Russian dolls — referred to as multiwalled tubes. There is now evidence that Iijima was not the first to discover carbon nanotubes (see 'Who saw the first nanotube?', overleaf), but after his 1991 paper, researchers woke up to the tubes' potential for the first time.

Initially, the multiwalled tubes posed a



problem for theorists, who struggled to predict their physical and electrical properties. But the simpler structure of single-walled tubes was easier for calculations, and the predictions provided much of the impetus for subsequent research. According to the theorists, single-walled tubes would be extremely stiff and would behave as either metallic conductors or as semiconductors, depending on how the rows of hexagons were orientated relative to the tube's length. Multiwalled tubes were thought to contain a mixture of both types, but their single-walled counterparts offered the chance to study the 'pure' electrical behaviour.

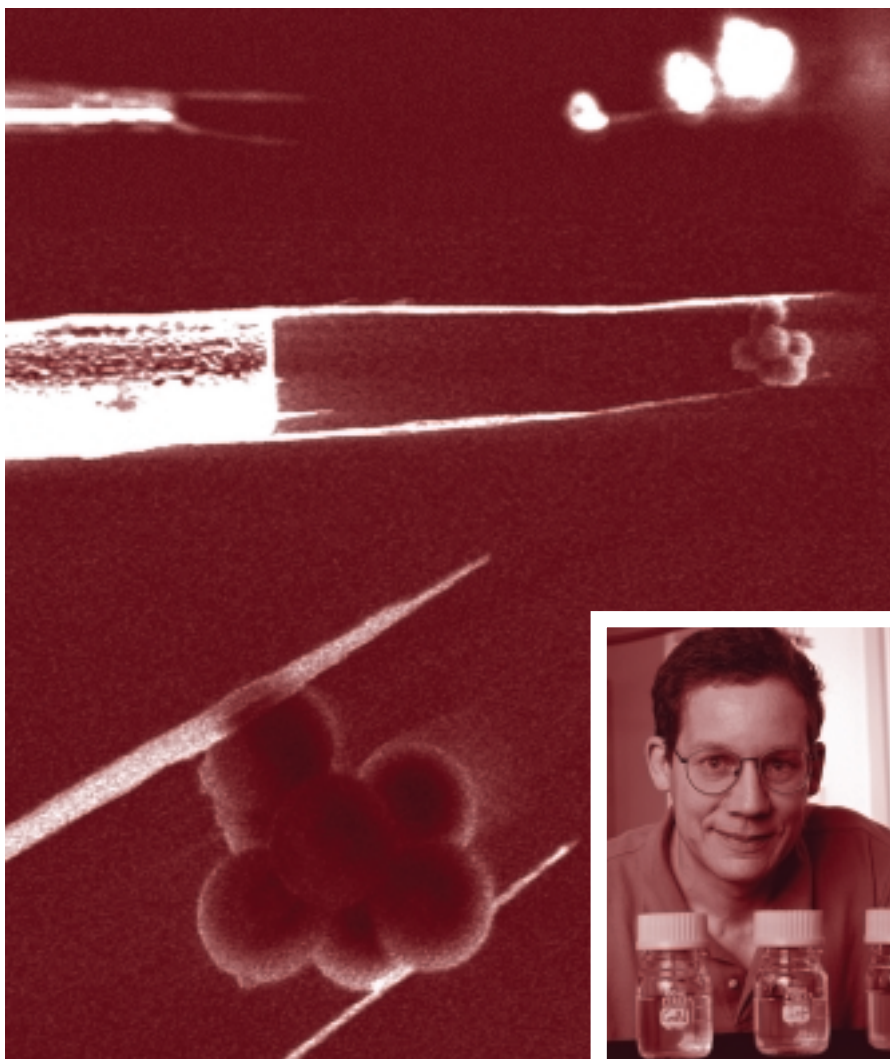
But it wasn't until 1993 that researchers were able to make such tubes. Iijima's team<sup>8</sup> and Donald Bethune's group at IBM's Almaden Research Center in San Jose, California<sup>9</sup>, simultaneously discovered that the growth of single-walled nanotubes is promoted by mixing cobalt or iron into the graphite electrodes.

### Carbon connections

Electrical conductivity was perhaps the tubes' most enticing property. Commercial techniques for making silicon chips can currently create wires about 180 nanometres wide. If nanotubes could be used as wires, this width would be reduced by at least an order of magnitude. In addition, the promise that some tubes would behave as semiconductors led researchers to suspect that they could be used to form nanoscale versions of components such as transistors — electronic devices that can switch or amplify current. This year, nanotubes have even been shown to behave as superconductors, materials that lose all of their electrical resistance below a certain transition temperature. So far, the highest transition temperature for nanotubes is 15 K (ref. 10), which is too low for practical applications.

Hongjie Dai and his colleagues at Stanford University took a step towards electronic circuits based on nanotubes when, in 1998, they showed how individual single-walled nanotubes might be grown directly onto a simple electrical circuit<sup>11</sup>. Dai's nanotubes formed between metallic islands on a silicon surface as heat was used to break down the methane gas surrounding the surface.

With modification, nanotubes can also become electrical devices in their own right. In 1998, Cees Dekker's group at the Delft University of Technology in the Netherlands turned a nanotube into a transistor<sup>12</sup>. Dekker passed current through the nanotube using metal electrodes attached to each end of the tube. A voltage applied to a third electrode underneath the tube altered the way in which the nanotube conducted electricity, and could be used to switch current flowing through the nanotube on and off. Other groups have since produced different types of nanotube transistor.



In a pinch: Charles Lieber (inset) has made nanotube 'tweezers', shown here picking up 350-nm beads.

Transistors are the basic components in logic gates, the electronic devices used to build computers. This August, Phaedon Avouris and his colleagues at IBM's Thomas J. Watson Research Center in Yorktown, New York, described how, by wiring two types of nanotube transistors together, they created the simplest type of logic gate — one that inverts its input, turning a binary '1' into a '0' and a '0' into a '1' (ref. 4). Dekker's group, meanwhile, has made gates with other, more complex behaviour<sup>5</sup>.

### Uphill struggle

Despite the promise of this work, conventional silicon technologies are so well established that replacing them will be difficult. "I would not expect to see the commercial use of nanotubes in logic or computer applications within this decade," Dekker says. In addition, cheap production of nanotube-based circuits might depend on other technologies, such as molecular self-assembly<sup>13</sup>, which have yet to mature.

Researchers are also getting excited about the mechanical properties of nanotubes. Single-walled tubes are very strong, and are

stiffer than the conventional carbon fibres currently used in composite materials. They can also be bent double and yet spring back into shape without suffering any apparent damage.

Such properties have led to some surprising applications. For example, the resolution of scanning probe microscopes — which provide images of a surface by moving a fine tip across it — depends on the sharpness of the probe's tip. And the longer the tip, the more deeply the microscope can probe into surface cavities. In 1996, Smalley and colleagues showed that mounting a nanotube on the microscope's tip improves both resolution and penetration<sup>2</sup> — and the tube's strength means that it will simply bend, instead of snapping, if it is pushed a little too far into the surface. Charles Lieber's group at Harvard University has since extended the concept by adding chemical groups to the end of the tips, so that the microscope can recognize the chemical composition of the surface being studied<sup>14</sup>. Further work indicates that a wide range of different substances, including biological molecules, can be attached to the tubes<sup>15,16</sup>.



## news feature

Together with his colleague Philip Kim, who is now at Columbia University in New York, Lieber has also used nanotubes as extremely fine tweezers<sup>3</sup>. They attached two nanotubes to a glass rod patterned with two electrodes. Giving these electrodes different charges creates an electrostatic attraction between the nanotubes, which makes them bend towards each other. This allowed Lieber and Kim to use the tubes to pick up and move tiny polymer beads. A team led by Yoshikazu Nakayama at Osaka Prefecture University in Japan has recently created similar tweezers on the tip of a scanning probe microscope<sup>17</sup>, raising the possibility that the microscope could be used both to observe and to manipulate samples, such as individual cells.

Exciting as these projects are, all exploit the physical properties of nanotubes at the microscopic scale. Materials scientists would like to take advantage of these properties over longer length scales, but at present, nanotubes cannot be grown longer than a millimetre or so. Last year, a team led by Philippe Poulin of the University of Bordeaux in France created fibres and ribbons from a tangled web of nanotubes<sup>18</sup>; but the tubes were not continuous, so the ribbons lacked the strength of individual nanotubes.

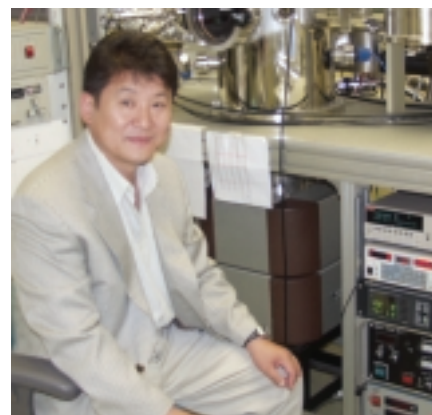
Nevertheless, commercial applications are

emerging. In 1999, Jong-min Kim and his colleagues at the Samsung Advanced Institute of Technology in Suwon, Korea, developed a flat-screen colour display in which nanotubes supply the electron beams that cause phosphors on the screen to light up<sup>19</sup>.

### Screen shots

The nanotubes act as a crucial component of miniature cathode-ray tubes — the bulky devices used in conventional televisions. Strong electric fields are used to pull electrons from the tips of the nanotubes, and the electrons are focused into a beam by tiny electrodes. Similar devices have been under development for many years, using electron emitters made from semiconductors or diamond. But nanotubes have sharper ends than the tips that can be made from these other materials — and the sharper the emitter's tip, the lower the electric field needed to extract electrons.

For such applications to be commercially viable, nanotubes will have to be readily available and inexpensive. Last year Smalley co-founded Carbon Nanotechnologies Incorporated (CNI), a Houston-based company geared to producing nanotubes. This September, CNI set to work on a new manufacturing system to produce 200 grams



**Jong-min Kim believes nanotubes could replace cathode-ray tubes to make flat TV screens.**

of nanotubes per day. At around US\$500 per gram, the tubes are still not cheap, but mass production should drive costs down. CNI hopes to make around 9 kilograms a day by 2002, and could be turning out thousands of kilograms per week by 2004. "In time, millions of tonnes of nanotubes will be produced worldwide every year," predicts Smalley.

As applications and commercial uses for nanotubes begin to emerge, demand is sure to rise. And research in the field continues unabated — futuristic applications, including fuel tanks for future hydrogen-powered vehicles<sup>20</sup> are under investigation. Some may well fall by the wayside, but it now seems inconceivable that Iijima's discovery will not make its impact felt beyond the laboratory. ■

**Philip Ball is a consultant editor for *Nature*.**

## Who saw the first nanotube?



The history of carbon nanotubes may be a little longer than was previously thought. Richard Smalley of Rice University in Houston, Texas, points out that nanotubes could have been unknowingly produced by late nineteenth century chemists experimenting on methane<sup>21</sup>. And in the 1960s and 1970s, at least two groups made and characterized the tubes, only for their discoveries to go largely unnoticed.

Roger Bacon of the National Carbon Company in Parma, Ohio, then part of Union Carbide, produced nanoscale scrolls of graphite in 1960 (ref. 22), confirming their dimensions and structure using microscopy and X-ray diffraction. Bacon's work may have been ignored by science historians, says Smalley, because of the view that he made scrolls. But recent work<sup>23</sup> has shown that Sumio Iijima at NEC in Tsukuba, Japan, who in 1991 published what is often thought to be the first paper on nanotubes, was using a procedure that generates a mixture of scrolls and tubes, suggesting that Bacon may also have done this.



In the late 1970s, Peter Wiles (bottom, left) and John Abrahamson (top, left) of the University of Canterbury in Christchurch, New Zealand, went one step further. They were investigating the carbon fibres produced when sparks were passed between two graphite electrodes, and found that one of the electrodes was coated with "mats of small fibres"<sup>24</sup>. In 1979, electron-diffraction measurements showed that the walls of these 'fibres' were made of carbon in a graphite-like arrangement. They described the tubes as made of several layers of crystalline carbon "wrapped together" — they had discovered multiwalled carbon nanotubes.

Wiles and Abrahamson presented their findings at the 14th Biennial Conference on Carbon, held at Pennsylvania State University in June 1979.

An extended abstract of that presentation was published at the time<sup>25</sup>, but no further reports appeared in print.

The contrast with the excited reception of Iijima's paper is remarkable, and illustrates how new discoveries rely on the prevailing research climate. In an age before nanotechnology and the fullerene-induced interest in carbon chemistry, there seemed little reason to regard the fibres as anything more than a smaller versions of familiar micrometre-scale carbon fibres. "Although we figured we had found something novel, we never for a moment imagined the eventual significance of the discovery," says Wiles.

1. Iijima, S. *Nature* **354**, 56–58 (1991).
2. Dai, H., Hafner, J. H., Rinzler, A. G., Colbert, D. T. & Smalley, R. E. *Nature* **384**, 147–150 (1996).
3. Kim, P. & Lieber, C. M. *Science* **286**, 2148–2150 (1999).
4. Derycke, V., Martel, R., Appenzeller, J. & Avouris, Ph. *Nano Lett.* **1**, 453–456 (2001).
5. Bachtold, A., Hadley, P., Nakanishi, T. & Dekker, C. *Science* **4** October 2001 (10.1126/science.1065824).
6. Kroto, H. W., Heath, J. R., O'Brien, S. C., Curl, R. F. & Smalley, R. E. *Nature* **318**, 162–163 (1985).
7. Krätschmer, W., Lamb, L. D., Fostiropoulos K. & Huffman, D. R. *Nature* **347**, 354–358 (1990).
8. Iijima, S. & Ichihashi, T. *Nature* **363**, 603–605 (1993).
9. Bethune, D. S. *et al.* *Nature* **363**, 605–607 (1993).
10. Tang, Z. K. *et al.* *Science* **292**, 2462–2465 (2001).
11. Kong, J., Soh, H. T., Cassell, A. M., Quate, C. F. & Dai, H., *Nature* **395**, 878–881 (1998).
12. Tans, S. J., Verschueren, A. R. M. & Dekker, C. *Nature* **393**, 49–52 (1998).
13. Ball, P. *Nature* **413**, 667–668 (2001).
14. Wong, S. S., Joselevich, E., Woolley, A. T., Cheung, C. L. & Lieber, C. M. *Nature* **394**, 52–55 (1998).
15. Erlanger, B. F., Chen, B.-X., Zhu, M. & Brus, L. *Nano Lett.* **1**, 465–467 (2001).
16. Star, A. *et al.* *Angew. Chem. Int. Ed.* **40**, 1721–1725 (2001).
17. Akita, S. *et al.* *Appl. Phys. Lett.* **79**, 1691–1693 (2001).
18. Vigolo, B. *et al.* *Science* **290**, 1331–1334 (2000).
19. Choi, W. B. *et al.* *Appl. Phys. Lett.* **75**, 3129–3131 (1999).
20. Zandonella, C. *Nature* **410**, 734–735 (2001).
21. Hughes, T. V. & Chambers, C. R. US Patent 405,480 (1889).
22. Bacon, R. *J. Appl. Phys.* **31**, 283 (1960).
23. Mordkovich, V. Z. *et al.* in *Supercarbon: Synthesis, Properties, and Applications* (eds Yoshimura, S. & Chang, R. P. H.) 107–118 (Springer, Berlin, 1998).
24. Wiles, P. G. & Abrahamson, J. *Carbon* **16**, 341–349 (1978).
25. Abrahamson, J., Wiles, P. G. & Rhoades, B. L. Abstract in proceedings of the 14th Biennial Conf. On Carbon, Pennsylvania State Univ., June 1979. Reprinted in *Carbon* **37**, 1873–1874 (1999).

# Young, gifted...and spurned

While their contemporaries in other countries zip from postdoc to postdoc, young French scientists struggle to find work. Sally Goodman reports.

**I**'ll just have to find any old job in a supermarket if I'm going to continue to feed my children," says geneticist Claire Amadou. For the past 18 months, Amadou has been working for free in a lab in Toulouse run by the CNRS, France's national research agency, living off her husband's salary. Having just discovered that the latest in a long line of applications for funding has failed, she is now at the end of her tether. "All the other labs I've contacted have their own postdocs in the same situation. It's hopeless."

Amadou's plight is all too common in France. Scientists in universities and government labs are classed as civil servants — a protected breed who cannot be employed on temporary contracts. Limited numbers of postdoc contracts are funded by research charities and industry. But with no official postdoc system, the 7,000 PhD scientists the country churns out each year have few options. A lucky few land a coveted permanent job in the public research system; others find places in industry. But many more are left scrabbling around for scraps of funding, or find themselves out of work.

The government recognizes that there is a problem. Given that 40% of research staff in public labs and universities will retire over the next 10 years, officials realize that they cannot afford to let a generation of talent go to waste. To prevent their loss, science minister Roger-Gérard Schwartzberg plans to create 800 new permanent posts in the public research agencies between now and 2010. At a press conference in Paris on 24 October, he talked of "an historic occasion to rejuvenate French research".

Schwartzberg's plan — together with a

parallel effort announced in November 2000 to open 2,600 positions in the universities, a majority of which are likely to be in scientific disciplines — will provide a lifeline for many young French scientists. But critics argue that it leaves the underlying problems untouched. In essence, it is trying to fix the sclerosis caused by a system that has no stepping stone between a PhD and a permanent job by creating yet more permanent positions.

## Bizarre situations

The effective ban on universities and government research agencies employing French postdocs on short-term contracts creates some bizarre situations: advertisements for such positions often state that only foreign nationals will be considered. The fact that researchers must be given tenured posts from the beginning breeds conservatism and a lack of mobility between labs — hardly conducive to creating a vibrant research base. "Laboratories often prefer to recruit someone they know," says Rémi Barré, executive director of the public-interest group OST, which monitors trends in science and technology. "Anyone that moves off the defined path has a lot of problems finding a place." Amadou, for instance, spent three-and-a-half years working on a postdoc in Dallas, Texas.

Indeed, laboratories often design job descriptions to fit their own local candidates, and information on posts coming up is scarce. The recruitment process involves a laborious system of annual competitions, which is not well adapted to the needs of fast-changing research organizations. "The restrictions in place are extremely prohibitive" says Geneviève Berger, director general of the CNRS, "which is incomprehensible considering the number of potential recruits available and the budgets available to laboratories to fund them."

Some agencies are creating their own *de facto* postdoc programmes — although they are careful to describe payments as grants, not salaries, to avoid openly defying the rules. The medical research agency INSERM will in 2002 award 25 postdoc grants under an elite programme called Avenir — meaning 'future'. Researchers who are taken on will be given lab space and a working budget over three years with, at the end, the possibility of a permanent post. They will also be paid FF15,000 (US\$2,000) a month — about 50% more than most junior



Gainfully employed: but many young French scientists find it impossible to get a suitable job.

researchers. Similar moves are afoot in leading independent medical research centres. The Pasteur Institute in Paris has instituted a scheme that will employ researchers on a five-year trial basis, following a model already practised at the Curie Institute, also in Paris.

## French resistance

But across most of the public and academic research system, inflexibility remains the norm. "Recruiting someone for life at 35 means that you may realize after five years that you have a super technician, but someone without the intellectual capacity to create their own research environment," says Daniel Louvard, director of research at the Curie Institute.

Solving that problem will require a brave science minister. Creating a genuine postdoc system would mean negotiating an exception from the rules governing the employment of civil servants. This system is protected by powerful vested interests, which will fight any exceptions tooth and nail. Science trade unions also oppose the introduction of postdoc contracts.

In the meantime, Schwartzberg's plan is the best that is on offer. But many young scientists fear it will not be enough. As Amadou puts it: "Why would anyone want to be a researcher in France?" ■

Sally Goodman is *Nature's* French correspondent.

RICHARD LAMOUREUX/CNRS PHOTO THEQUE

AFP PHOTO JACK GUEZ



Roger-Gérard Schwartzberg's plan to create new permanent posts misses the point, say critics.

# BSE fostered by cosy relationships and lack of independent advice

Japan needs experienced scientists in positions of power to make reliable assessments of government advice

*Sir*— Your Opinion article “Japan’s beef scandal” (*Nature* **413**, 333; 2001) had an extraordinary impact on the Japanese mass media. Many newspapers and television news shows criticized the government’s irresponsible attitude to bovine spongiform encephalopathy (BSE, or mad-cow disease) by citing your article.

You pointed out the cosy relationships between government and industry and between government and healthcare institutions, strengthened by the former ministry officials, known as *amakudari* (because they have ‘descended from heaven’), who retired from government to join large companies. This is a major reason why the government did not take immediate action to prevent the spread of this dreadful disease. It seems to me that official comments about a “lack of scientific evidence” for the risk of BSE development results from a lack of scientists in the government.

Although many scientists with PhDs are now working for the government, most of them are in scientific research institutes belonging to the ministries. Only a few scientists with research experience are high-ranking officials. Political decisions related to science and health are made by recommendations from government advisory councils. Because high-ranking officials cannot understand the scientific

logic behind this advice, they tend to accept it without criticism or question. Therefore, political decisions are almost completely dependent on the people appointed to these advisory boards. Cosy relationships between government and industry are fostered when officials select members of advisory councils.

I believe that Japan should make a prompt effort to increase the number of scientists with appropriate research experience in the various ministries, especially those for health and science. Good scientists can well understand from their experience how trivial problems can lead to a fatal catastrophe. In addition, they have been trained to make judgements based on scientific logic, rather than to accept assertions uncritically.

I think that Japanese scientists and scientific societies should also feel some responsibility for the present situation. They have never spoken out or made efforts to increase the number of scientists in government. Nowadays, our daily lives cannot be detached from science. Japanese scientists should understand the present requirements of our nation, and take the initiative to change an irresponsible system.

**Mitsuo Tagaya**

*School of Life Science, Tokyo University of Pharmacy and Life Science, Hachioji 192-0392, Japan*

## Meat and bone meal still used in animal feed

*Sir*— In your News Story, “Japan’s first BSE case fuels fears elsewhere” (*Nature* **413**, 337; 2001), you state that “British farmers dumped [cattle feed containing animal parts] in the region [East Asia] when demand slumped in Europe in the early 1990s”. This is a careless statement, not up to the normal standards of your journal. British farmers did not and do not export animal feed. It is feed manufacturers who export feed.

As made clear in the exhaustive BSE inquiry report (see [www.bseinquiry.gov.uk](http://www.bseinquiry.gov.uk)), during 1989 and 1990 the UK authorities informed the Office International des Epizooties and the veterinary authorities in all countries importing meat and bone meal from the United Kingdom of the feed ban on UK domestic ruminants.

The use of meat and bone meal in feed for non-ruminants (pigs and poultry) was legal in the United Kingdom until August 1996, and until January 2001 in the rest of the European Union (EU).

Meat and bone meal is still being used in animal feed in many non-EU countries, despite the BSE experiences of the United Kingdom and of the rest of the EU.

**Stephen Rossides**

*Food, Health and Science Department, National Farmers’ Union, 164 Shaftesbury Avenue, London WC2H 8HL, UK*

## Tantalizing glimpse of a vanishing dinosaur

*Sir*— Your News item “Elusive fossil could conceal answer to dinosaur debate” (*Nature* **412**, 844, 2001) about the Chinese psittacosaurid with preserved integuments

prompts me to add some information. Early last year, I was invited to travel to the Museo Civico di Storia Naturale in Milan, and hence was able to examine the specimen in some detail. I was satisfied that the integumentary remains are real, and apparently different from the feathers or protofeathers seen on some theropod specimens from the same provenance.

At the time, my Italian colleagues in Milan hoped that a joint study of the specimen with Chinese colleagues could be arranged, after which the fossil would be returned to China. Together with my wife, Dr Haiyan Tong, and others, we tried to have the psittacosaur returned to a Chinese institution, by providing possible contacts in China. But our attempt failed and the specimen was removed from the Milan museum.

I do not know the exact reason for our failure, but one result was that plans for a joint study in Milan were dropped. I am ignorant of the present whereabouts of the specimen. Of course, I have always believed that the fossil should be returned to its country of origin, and I have no intention of publishing anything formal about it, other than this brief note, which I hope may shed some light.

**Eric Buffetaut**

*Centre National de la Recherche Scientifique, 16 cour du Liégat, 75013 Paris, France*

## Once again, insects worked it out first

*Sir*— Because of the contact between phases, transport of liquid droplets on solid surfaces is a complicated, non-efficient task requiring a considerable amount of energy. Increasing the contact angle between the liquid and solid (hydrophobicity) above 120° should ease the motion. By coating a small amount of liquid with grain spores of *Lycopodium*, Aussillous and Quéré<sup>1</sup> report their creation of non-sticky liquid marbles (contact angle of 180°). The droplets are soft solid and elastic, and can bounce and roll on a solid surface without leaking. It was predicted in the News and Views article<sup>2</sup> accompanying this paper that this brilliant trick will have a rapid practical use in wear-free micro-machines and lubrication.

Yet tiny, relatively primitive insects faced and solved a similar problem millions of years ago. Homopterans (aphids, whiteflies, and so on) feed on plant sap (usually phloem) that contains a high concentration of sugars but low nitrogen levels. Therefore, the extra sugars are deposited by the insects as sticky droplets, commonly known as honeydew. Some aphids live in galls where they face



the danger of contamination and drowning in their own honeydew. These insects coat the honeydew with wax secreted from specialized glands<sup>3</sup>.

Furthermore, some gall-forming aphid species clean their galls by pushing the coated honeydew outside<sup>4–6</sup>. The coated elastic honeydew droplets can be pushed, rolled and squeezed without leaking or wetting. Smith<sup>3</sup> noted that this behaviour was initially observed by Buckton in 1876 and may have promoted the evolution of social organization in aphids<sup>7</sup>. Practically, it is worthwhile examining whether aphid's wax is better at coating liquid droplets than are the spores of *Lycopodium*.

In short, 'liquid marbles' are yet another example of how insects 'developed' a technology through natural selection long before humans got around to it<sup>6</sup>. Technologists (high and low) and engineers should look for new solutions with the eyes of biologists.

**M. Inbar\*, J. C. Schultz†**

\*Department of Biology, University of Haifa at Oranim, Tivon 36006, Israel

†Department of Entomology, Pennsylvania State University, University Park, Pennsylvania 16802, USA

1. Aussillous, P. & Quere, D. *Nature* **411**, 924–927 (2001).
2. Mahadevan, L. *Nature* **411**, 895–896 (2001).
3. Smith, R. G. *Journal of Natural History* **33**, 513–530 (1999).
4. Aoki, S. *Kontyû* **48**, 71–73 (1980).
5. Kurosu, U. & Aoki, S. *Journal of Ethology* **9**, 51–55 (1991).
6. Benton, T. G. & Foster, W. A. *Proceedings of the Royal Society of London* **247**, 199–202 (1992).
7. Stern, D. L. & Foster, W. A. *Biological Review* **71**, 27–79 (1996).
8. Akre, R. D. *et al.* *Insects Did it First* (Ye Galleon Press, Washington, 1992).

## Disclosure of interests: there's a long way to go

Sir — I read with great interest and approval your policy (*Nature* **412**, 751; 2001) requiring authors to disclose financial interests. This policy has been sorely needed for a long time, and I hope that other journals follow suit. My only caveat is with respect to your anticipation that "employers" will "police" the policy.

I recently attended a symposium chaired by two key university counsel from two leading US universities (one private, one state), and asked them what they and their institutions do to ensure that scientists disclose financial interests in publishing articles. The answer was a blank stare — they assumed that such interests are disclosed as appropriate and have no procedures or policies for ensuring that such disclosures occur.

**Tamsen Valoir**

Jenkins & Gilchrist, 1100 Louisiana Street, Suite 1800, Houston, Texas 77002, USA

## Sequenced strains must be saved from extinction

Sir — Many prokaryotic strains used for genome sequencing projects are poorly documented and not generally available.

With improvements in sequencing technology and growing recognition of the value of microbial genome sequence data, the number of microbial genome-sequencing projects is increasing rapidly. There are 56 completed prokaryotic genome sequences (10 strains of Archaea and 46 of Bacteria), and another 210 in progress (see, for example, [www.tigr.org](http://www.tigr.org) and [www.integratedgenomics.com](http://www.integratedgenomics.com)).

Of these 266 projects, some of which are performed on organisms not available in pure culture — some endosymbionts, for example — only 51 represent the type strain of the species. (A type strain is made up of living cultures of an organism descended from the nomenclatural type.) Of the rest, 138 represent non-type strains (a non-type strain is often selected only because it happens to be close at hand); 31 projects concern symbionts and environmental (uncultured) strains; 32 do not specify a strain; 14 represent prokaryotic species with invalid species names (validly named bacterial species are either on the 1980 approved lists of names, or validated after 1980 by taxonomic description in the *International Journal of Systematic and Evolutionary Microbiology*, or by validation in that journal: an invalid name has no standing in nomenclature and may be changed subsequently. Only 115 represent the type species of the genus (the *typus* of the genus included when the genus name was originally validly published) and only 123 are deposited in public culture collections. Mandatory deposition of the type strain of any validly described (culturable) prokaryotic species in a major public culture collection guarantees the availability of the strain and allows cross-referencing of published data.

When there were only a few projects, these taxonomic and preservation issues were not so evident. With the explosion in sequencing and the sequencing of multiple strains of a species (including *Escherichia coli*, *Staphylococcus aureus* and other major pathogens), questions of strain identity and safekeeping assume more importance. Deposition of strains in public collections with long-term funding is the only way to ensure their maintenance and their continuing availability to the scientific community. As things stand, it is a real possibility that a strain for which a wealth of genomic data has been generated may become "extinct" through loss of viability.

We propose that the following standards should be adopted by the entire

community. First, genome-sequencing project lists and databases should include the name of the strain sequenced and its associated culture collection accession number(s), as well as its origin. Second, the type strain of a species should be used for sequencing unless other factors make this inappropriate. Third, strains for which genome sequences have been, or are being, generated should be deposited in at least two major public biological resource centres, such as the American Type Culture Collection, the German Collection of Microorganisms and Cell Cultures, the Pasteur Institute Collection or the Japanese Collection of Microorganisms. **Naomi Ward\*, Jonathan Eisen\*, Claire Fraser\*, Erko Stackebrandt†**

\*The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20850, USA

†German Collection of Microorganisms and Cell Cultures, Mascheroder Weg 1b, D-38124 Braunschweig, Germany

## Taxonomists make a name for themselves

Sir — Many publications by taxonomists are not included in the *Science Citation Index (SCI)*, as pointed out in Correspondence by A. G. Valdecasas *et al.* (*Nature* **403**, 698; 2000) and by F. T. Krell (*Nature* **405**, 507–508, 2000). E. Garfield (*Nature* **413**, 107; 2001), on the other hand, pointed out some citation classics in taxonomy based on the *SCI*.

Your correspondents fail to mention the real reason most taxonomists are not included in the *SCI*: the way taxonomic research is cited means it does not feature in references. In referring to the name of a species, it is customary to state the name(s) of the author(s) after the species name, sometimes abbreviated and with the year of publication of the first description of the species. This citation method is ignored by the *SCI* and other citation indexes, which is why these scientists do not get the credit they deserve, even though each use of a species name in fact represents a citation.

A quick look in the ISI *Web of Science* 1988–2001 [[www.isinet.com/isi/products/citation/wos/index.html](http://www.isinet.com/isi/products/citation/wos/index.html)] shows that the most frequently cited scientists are the authors who named the bacterium *Escherichia coli* (108,262 citations), the yeast *Saccharomyces cerevisiae* (43,403) and the fruitfly *Drosophila melanogaster* (14,451). To put these numbers into perspective, Albert Einstein is cited 'only' 11,920 times.

**Gerard van der Velde**

Department of Animal Ecology and Ecophysiology, University of Nijmegen, Nijmegen, Toernooiveld 1, 6525 ED, The Netherlands

# No need to worry about the future

Environmentally, we are told, 'things are getting better'.

## The Skeptical Environmentalist: Measuring the Real State of the World

by Bjørn Lomborg  
Cambridge University Press: 2001, 515 pp.  
£47.50, £17.95

Stuart Pimm and Jeff Harvey

The subtitle gives the book away. It rehashes books such as Ronald Bailey's *The True State of the Planet* (Free Press, 1995). As Bjørn Lomborg tells us, the book's origin was a class he taught in 1997. The original Danish version appeared a mere year later — remarkably fast, given the delays of academic publishing. It shows, too. This survey of global environmental problems — food, forests, energy, water, pollution, biodiversity, global warming — reads like a compilation of term papers from one of those classes from hell where one has to fail all the students. It is a mass of poorly digested material, deeply flawed in its selection of examples and analysis.

Lomborg admires the late Julian Simon, author of *The Ultimate Resource* (Princeton University Press, 1996). Beside Simon, Voltaire's optimistic Dr Pangloss is gloomy and Albert Einstein a theoretical novice. Simon impressed the US political right by his assertion that we have "the technology to feed an ever-growing population for the next 7 billion years". Ecologists were challenged by this remarkable rejection of basic ecological laws. At present growth rates, the human mass would exceed that of the biosphere within the millennium. Physicists should be in awe, too. Well before the allotted time, human mass would be expanding faster than the Universe.

Thus influenced, Lomborg begins with "the litany" — the list of things wrong with the planet, and why, when we see things his way, "things are getting better". The litany quotes news magazines and a book by two science-fiction writers, but not scientists directly. No external references support the ensuing paragraphs justifying that 'things are getting better'. Quoting the primary literature troubled Simon, too.

Like bad term papers, Lomborg's text relies heavily on secondary sources. Out of around 2,000 references, about 5% come from news sources and 30% from web downloads — readily accessible, therefore, but frequently not peer reviewed. A mere 1% are original papers in *Nature*, half as many again come from contributors to Simon's books. This bias towards non-peer-reviewed material over internationally reputable journals is sometimes incredible

— for example, the claim that the evidence for pollution at New York's Love Canal was "jaded". At other times it seems fictional. "Scientific luminaries such as Harvard biologist E. O. Wilson and Stanford biologist Paul Ehrlich are the enthusiastic supporters of an ambitious plan ... to move the entire population of the US. ... people would live in small enclosed city islands." The reference is directly attributable neither to Wilson nor to Ehrlich. "Is it true?" we asked them. Ehrlich: "I know of no such plan. If there were one, I wouldn't support it." Wilson concurred.

Lomborg's great optimism about humanity's future shows up in the way he presents statistics. In the hell-hole that is so much of sub-Saharan Africa, "starving people" constituted "38 percent in 1970 ... [but only] "33 percent ... in 1996. [The percentage is] expected to fall even further to 30 percent in 2010." The absolute numbers of starving are curiously missing from these paragraphs. Roughly, the region's population doubled between 1970 and 1996. To keep the numbers of starving constant, the percentage would have had to have dropped by more than half. The absolute numbers of

malnourished in the region — as well as those whom fate will spare through their death from the myriad consequences of poverty (including AIDS) — are surely inconsistent with the first-listed "global trend" in a chapter entitled "Things are getting better".

Often, Lomborg misses the critical literature in exactly the same ways as did Simon. For example, consider the chapter on biodiversity. It starts out with the by-now standard denigration of consensus estimates on extinction rates and omits relevant papers in even obvious places — including the paper demonstrating that Simon's estimates are three to four orders of magnitude below everyone else's.

The text employs the strategy of those who, for example, argue that gay men aren't dying of AIDS, that Jews weren't singled out by the Nazis for extermination, and so on. "Name those who have died!" demands a hypothetical critic, who then scorns the discrepancy between those few we know by name and the unnamed millions we infer. Exactly repeating Simon, Lomborg juxtaposes the small number of named dead species against the huge number of species

## Those with no future

The long, curved beak of the huia (*Heteralocha acutirostris*) distinguishes the female of this species of lost New Zealand bird from the shorter-beaked male. The bird was last sighted in 1907, and its distinctive tail feathers, prized by the Maori and by European trophy hunters, probably hastened its demise. *Extinct Birds* (new edition) by Errol Fuller (Oxford University Press/Cornell University Press, £29.50/\$49.95) tells the huia's story, together with those of other recently extinct bird species.



## book reviews

for which we have no knowledge at all. After pages of confused argument, his extinction estimate of "0.7 percent over the next 50 years" is strikingly discordant with the 10–40% of well-known species that teeter on the brink of extinction just from human actions to date. About 2% of well-known species are already so desperately rare that we don't know whether they do survive. Lomborg finds comfort when some are rediscovered. Like terminally ailing humans, their lingering survival does not allay fears about the unfolding epidemic.

On future trends based on forest losses, his flawed examples are unoriginal. "In the US, the eastern forests were reduced ... to fragments totalling just 1–2% of the original area ... this resulted in the extinction of only one forest bird". The correct percentage is close to 50%, and the number of extinctions four, plus two seriously wounded. Those extinctions constitute 15% of the bird species found only within the region (the only ones at risk of global extinction). They strikingly confirm the predictions made from the species-area models that Lomborg disparages.

An industry has arisen debunking this book chapter by chapter. At present, it includes a website ([www.anti-lomborg.com](http://www.anti-lomborg.com)); a series of essays planned for *Scientific American*; a guide for journalists documenting Lomborg's more egregious errors being assembled by the Union of Concerned Scientists; and various published pamphlets. We have provided only a sampler.

But *Nature* instructs its reviewers to do more than merely describe a book's contents; we must examine its wider implications. The only such implication we see causes us to ask why Cambridge University Press would decide to publish a hastily prepared book on complex scientific issues which disagrees with the broad scientific consensus, using arguments too often supported by news sources rather than by peer-reviewed publications. Certainly, controversy is part of science, but extraordinary claims require the extraordinary scrutiny that comes from competent peer review — something that appears to be missing in this case. ■

Stuart Pimm is at the Center for Environmental Research and Conservation, MC 5556, Columbia University, 1200 Amsterdam Avenue, New York, New York 10027, USA. Jeff Harvey is at the Centre for Terrestrial Ecology, Netherlands Institute of Ecology, PO Box 40, 6666 ZG Heteren, The Netherlands.

### More on environmental risk analysis The Precautionary Principle: A Critical Appraisal of Environmental Risk Assessment

by Indur M. Goklany  
Cato Institute, \$17.95



Bleak prospect: Scott's diligent study of the polar weather could not prepare him for what lay ahead.

## Tragic outcome of extreme conditions

### The Coldest March: Scott's Fatal Antarctic Expedition

By Susan Solomon  
Yale University Press: 2001. 416 pp.  
\$29.95, £19.95

Cornelia Lüdecke

Many books have been written about Robert Falcon Scott's final, ill-fated expedition to Antarctica in 1910–12, where he died during the return trip from the South Pole, only a few miles away from the nearest safe depot. It is well known that Scott relied on motor sledges, used weak ponies and allowed five men to go south instead of the planned four. These facts have contributed to the legend that Scott's failure was a result of ineptitude. Susan Solomon, an experienced polar researcher and leader of the American National Ozone Expedition, shines new light on Scott's tragic leadership error in not being able to predict the unexpected — unlike Roald Amundsen, who won the race. She analyses both published and unpublished material to show why Scott died — disproving the criticisms of the legend. An important aspect of her analysis is the emphasis she places on data provided by automatic weather stations installed close to Scott's route, which have been used since 1985 to facilitate flights to the scientific station at the South Pole.

Each chapter is introduced by modern reflections of a 'visitor' strolling around the Pole Station, whose experiences are compared with those of Scott; the identified problems are addressed in the ensuing text. The first three-quarters of the book follow Scott's expeditions chronologically. We are told what his experiences taught him during his first expedition to Antarctica between

1901 and 1904, and about his preparations for the second — his transport planning, and his questionable decisions. But the text focuses on the extremely cold temperatures experienced by Edward Wilson and his two comrades on the Windless Bight south of Ross Island, and, even more so, on the bad weather Scott's group encountered during their trek to the pole. We follow the group of five beyond the "H of Hell" to the "awful place", shivering with them from blizzards and very low temperatures, and sharing their hunger and thirst. The human story, which has been pieced together from the accounts of the men involved, is illuminating.

Modern weather data are used in a new way to provide insight into the misfortune of the fatal journey. Scott was aware that meteorology would play a major part in this polar trek. He thoroughly investigated weather conditions during various trips in March, July and September 1911, and discovered that temperatures at the barrier (the Ross Ice Shelf) were about 20 °F lower than at the Cape Evens base camp. A minimum thermometer at One Ton Camp (on the barrier) recorded a temperature of –72 °F in 1911, whereas measurements at Cape Evens revealed a long, 'coreless' winter with mean temperatures of about –23 °F. George C. Simpson, the expedition's meteorologist, predicted that Scott would have to face a short and cold summer during his trek, with temperatures rising by 20–30 degrees during blizzards and falling afterwards.

Unfortunately, Scott decided to start late in 1911, because of the poor condition of his ponies, which should have pulled the equipment. After the ponies collapsed, man-hauling the heavy sledges became an exhausting task. Despite this, they arrived at the pole on 17 January 1912. On their way back, very cold temperatures caused the snow surface to become like sandpaper — the sledges could barely be moved, because there was no further melting of ice crystals

FROM THE VOYAGE OF THE DISCOVERY VOL. I



for lubrication. The ice plateau (nearly 10,000 feet high) also took its toll. The men lost a great deal of weight and suffered from dehydration, which was caused by a shortage of oil for preparing hot meals and melting the ice. Their pace slowed, they endured severe frostbite and needed much more time than anticipated to cover the distances between depots.

The last part of the book adds more detailed information from historical and modern weather data, as well as diaries and letters. Solomon shows that Scott's party encountered a very unusual cold snap at the end of February 1912, unlike any recorded since. She critically evaluates all the data taken during the crucial months, and compares these with modern temperature data. The clue to her analysis is provided by katabatic winds, which originate in the cold, heavy air that flows down from the ice cap. Rushing down the steep slopes to the coastline, these winds warm and develop into blizzards. Once the air has left the pole, it takes some time to build up another reservoir, and consequently these blizzards cease after a couple of days. Clearly, Scott could not have experienced a 10-day blizzard. Solomon's hypothesis is that Scott could have made it to the next depot if the weather had not been extremely cold beforehand; he was then doubly unlucky in being forced by frostbite to stay in their tent even after the blizzard was over. Here he died, together with his comrades Henry Bowers and Edward Wilson, who dared not leave him alone.

*The Coldest March* is well researched and

well written, and should appeal to a broad readership, as well as to meteorologists and polar historians. ■

Cornelia Lüdecke is at the Institute of History of Natural Sciences, Ludwig Maximilians University, Museumsinsel, D-80306 Munich, Germany.

#### More on polar exploration

#### **The Voyage of the *Discovery*: Scott's First Antarctic Expedition, Vols I & II**

by Captain Robert F. Scott  
Cooper Square Press, \$35

#### **A Fabulous Kingdom: The Exploration of the Arctic**

by Charles Officer & Jake Page  
Oxford University Press, £17.99, \$25

## Bumps on the brain

#### **The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain**

by William R. Uttal  
MIT Press; 2001. 255 pp. \$39.95, £27.50

John C. Marshall

Phrenology is the vulgar Victorian term that refers to a set of conjectures first advanced by the Viennese physician Franz Joseph Gall (1758–1828). In addition to his exemplary clinical skills, Gall was the greatest neuro-anatomist of his day, and he also set the agenda for the subsequent development of

cognitive neuroscience. In brief, Gall argued that, just as the body from the neck down is comprised of distinct organs, each associated with particular physiological functions, so likewise is the body from the neck up. According to this hypothesis, the brain itself consists of many 'mental organs', each dedicated to a particular cognitive, emotional or voluntary function.

As William Uttal writes in the first chapter of *The New Phrenology*, the gallerist paradigm thus asks three questions: "1. Can the mind be subdivided into components, modules, or parts? 2. Does the brain operate as an equipotential mass, or is it divisible into interacting but separable functional units? 3. Can these components—assuming they exist in some valid psychobiological sense—be assigned to localized portions of the brain?"

If these questions are interpreted fairly broadly, there can be few neuroscientists who would not today answer yes on all counts. But, as usual, the devil is in the detail. And it is those particulars that Uttal claims to examine in this highly polemical tract. Most interestingly, Uttal seems to be one of the few neuroscientists (his background is in electrophysiology) who are inclined to plead that the localization of cognitive processes in the brain is a chimera. Or, more precisely, he contends that there are so many "limits and constraints" on positive answers to any of the gallerist questions that the reply might as well be no.

It is, of course, true that many different taxonomies of cognitive processes have been proposed over the past 2,000 years. And it is also true, as Uttal points out, that most of these classifications are pretty silly. Yet one might have expected that anyone concerned with the psychobiological categorization of psychic functions would have seriously considered the mental modules deployed by late-nineteenth-century neurologists to classify the cognitive disorders that could result from relatively focal damage to the human brain. The capacity to recognize objects in one's environment is compromised in the agnosias; the ability to assess where those objects are located is perturbed in a wide range of topographic disorders; the performance of skilled actions is impaired in the apraxias; remembrance of the past is clouded in the retrograde amnesias; the ability to communicate is weakened in the aphasias. In all these instances, careful clinical studies have established that different components of each of these mental modules can be independently impaired after brain damage.

Detailed neuropathology at autopsy (and, more recently, the *in vivo* visualization of structural lesions by such techniques as magnetic resonance imaging) then showed that distinct lesion sites could often be associated with different types of

## Medical masterpieces

As an excellent excuse for displaying sumptuous works of art, *The Journal of the American Medical Association* began in the 1960s to put a reproduction of fine art in place of the Table of Contents that had figured on its cover since the journal started in the 1880s. For more than 25 years an essay has accompanied the image—to show, as the author of most of the essays says, "that everything is related: art, medicine, life". *The Art of JAMA II: Covers and Essays From The Journal of the American Medical Association* by M. Therese Southgate (AMA Press, \$69.95) contains a selection of covers, including the Goya painting *Bartholomé Sureda y Miserol* shown here, and their accompanying essays.



## book reviews

impairment to each of the cognitive sub-domains in question.

Uttal is well aware that “the brain is not homogenous but made up of regions that are differentiated by function”. And he is equally aware of some of the clinico-anatomical correlations that have been established in behavioural neurology. But he then proceeds to give “a totally different conceptualization of the localization problem”. This restatement “offers, in place of a specific function being precisely localized (that is, instantiated, represented or encoded) in a particular place, the idea of one centre contributing to the operation of a complex system of nodes and loci that are collectively responsible for the behaviour.” Well, yes: this is precisely what every behavioural neurologist and neuropsychologist has argued since (at least) Carl Wernicke’s fractionation of the ‘aphasic symptom complex’ in 1874.

Sadly, Uttal does not engage with the details of this literature, despite acknowledging how clinical evidence supports the claim that complex cognitive processes such as speech and calculation “involve widely dispersed regions of the cerebral cortex”. Rather, he rushes down the sexier path of functional neuroimaging to discuss such techniques as electroencephalography, positron-emission tomography and functional magnetic resonance imaging. Here he draws attention to failures of replication in which seemingly slight changes in experimental design, stimulus materials and task characteristics can lead to very different patterns of brain activation.

These are all genuine problems, but what needs to be debated here is the extent to which they reflect our failure to design good experiments rather than the brain’s failure to instantiate distributed localization of function. Uttal also notes, correctly, that individual differences in brain organization can undermine the reliability of associations between structure and function. Again, however, this is a challenge to discover the range and limits of variation rather than a cause for despair. A judicious conjunction of psychophysical experimentation, functional neuroimaging, and lesion studies will no doubt eventually elucidate the basic functional architecture of cognition. We can then start on the really difficult issue of showing how, not where, neuroanatomical and neurochemical structures and processes instantiate our best guesses about the organization of the mind.

Many of the problems of functional localization that Uttal outlines are real enough, but he offers no new ways of solving them and no convincing arguments that some other approach would resolve them. ■

John C. Marshall is in the University Department of Clinical Neurology, Radcliffe Infirmary, Woodstock Road, Oxford OX2 6HE, UK.

## Science in culture

### Rocket retrospective

*Frau im Mond*, directed by Fritz Lang, an eerily predictive tale of a journey to the Moon.

Heike Langenberg

A slender rocket, destined for the Moon, appears on the screen. Accompanied by cheering masses and ceremonial speeches, it makes its way from the hangar to the floating launch station. The tension builds, the countdown runs, and the rocket takes off into space.

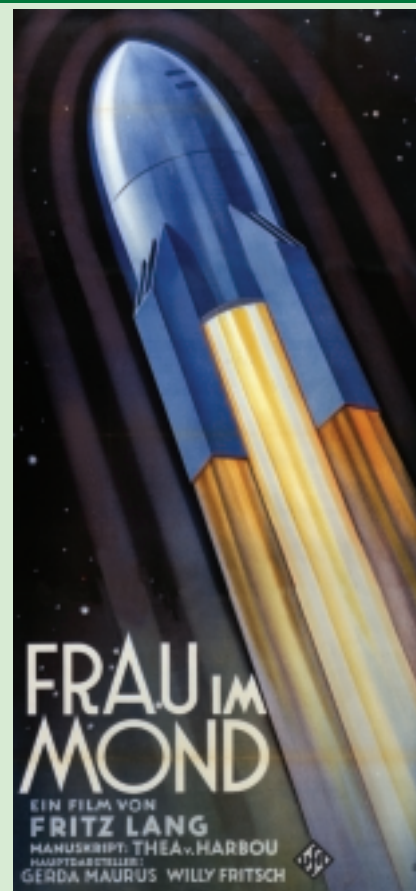
This is not 16 July 1969, when Apollo 11 lifted off for the Moon. It is part of Fritz Lang’s last silent movie, *Frau im Mond* (“Woman in the Moon”), released in 1929 and long confined to the archives. The film has now been reconstructed from the original negatives in a cooperation between the German Federal Film Archive and the Friedrich-Wilhelm-Murnau Foundation in Wiesbaden, and was first presented in superb quality at this year’s international film festival in Berlin.

Back in 1929, German film critics were not impressed by the pompous opening night of *Frau im Mond*. The *Berliner Börsen-Courier* newspaper declared that Lang’s talents were wasted on his wife’s scripts, and the *Vossische Zeitung* dismissed the film as a movie for 14-year-old boys. Nevertheless, *Frau im Mond* became the year’s most successful film in terms of sales. And it caught the imagination of an unusual 17-year-old, Wernher von Braun, who was to become head of German rocket development during the Second World War and later one of the leading scientists in the US programme to develop the powerful Saturn rockets. It was a Saturn V rocket that lifted Apollo 11 to the Moon.

Von Braun was already interested in rocketry. In 1925 he had read Hermann Oberth’s 1923 book *Die Rakete zu den Planetenräumen* (“The Rocket into Interplanetary Space”). The book allegedly spurred young Wernher’s interest in mathematics, and shot him to the top of his class.

Von Braun was not alone in his enthusiasm for Oberth’s rockets. Rocket fever gripped the whole of 1920s Germany, and inspired Lang to make his film. Lang wanted his work to be based on the latest scientific knowledge, and so asked Oberth to act as technical adviser. The rockets created for the film sets are startlingly similar to later rocket design.

The plot of the film is a kitsch tale of love and adventure. The flight to the Moon is motivated by an old professor’s idea that gold is hidden in the lunar mountains, and the main participants in the journey — apart from the professor and his mouse — are three rocket engineers: two young men and Fräulein Friede, who is engaged to one of the engineers but loved by both. An 11-year-old stowaway and an evil representative of global finance compound the lunacy. Thea von Harbou’s film script may have had its weaknesses: the film is over-long at three hours. Nonetheless, the scenes featuring the rocket are magnificent — and, after all, the audience loved it.



Shortly after the debut of *Frau im Mond*, von Braun joined the German Society for Space Travel, and assisted Oberth in his research on liquid-fuel rockets. Oberth had been contracted by Lang and the film company UFA (Universum Film AG) to build a man-sized rocket to be launched into the stratosphere as a gimmick for the film’s opening night. But the work was not finished in time. It wasn’t until 1931 that the liquid-fuel rocket was finally ready to be patented.

Money for rocket research was scarce — until the German army discovered the potential of rocketry for less peaceful ends. By 1933, Captain Walter Dornberger of the German Reichswehr had realized that rockets were potential weapons whose development was not regulated under the Treaty of Versailles. Von Braun received a research grant from the armed services, and it was not long before rocketry became classified research.

With the military takeover, funds were no longer a problem. But the scientists had sold their souls to the Nazis and the dream of flying to the Moon was abandoned — until the cold war provided a new backdrop for space travel. As a poignant reminder of the early years of enthusiasm, the first successful German V2, launched on 3 October 1942 and developed by von Braun and his associates, carried an emblem symbolizing *Frau im Mond*. ■

Heike Langenberg is a physical sciences editor at Nature.



# The tracks of thought

In both science and technology, metaphors direct the way we think, reason and hypothesize.

**Douwe Draaisma**

Once the nucleotides adenine, thymine, guanine and cytosine had lent their initials to the elements of the genetic code, the rest came naturally. These four letters form the 'genetic alphabet'. Each group of three letters, or codon, represents a 'word', implying that there are 42424464 words in the 'genetic lexicon'. The process of copying DNA into RNA is referred to as 'transcription'; the step from RNA to protein synthesis is called 'translation'. Perhaps the most elegant metaphor is the palindrome, which refers to symmetrical DNA clones that read the same in both directions. In *Who Wrote the Book of Life?* (Stanford University Press, 2000), Lily Kay argues that modern genetic theory has been dominated by the metaphor of reading and writing and its associated connotations of codes and cryptography.

Genetics is not the only scientific field to have felt the lure of this metaphor. We also tend to describe the processes that occur in machines in terms of reading and writing. For a computer to process information, data must be 'machine-readable'. Storing new information in the place that was previously occupied by older information is called 'overwriting'. Information, whether it comprises digits, music or graphs, can be 'written across' from working memory to the hard disk. Even the images stored in a hologram are 'read out' by laser beams. In both science and technology, the processes of reading and writing seem to have a firm grip on our nomenclature, if not on our imagination.

Language is both our most important vehicle of symbolic communication and the dominant medium in which information is externally stored. But if we wish to understand the seductive attraction of reading and writing as scientific metaphors, we should also be sensitive to the fact that

**M**etaphors, particularly the powerful ones such as reading and writing, lay the tracks for our trains of associations.

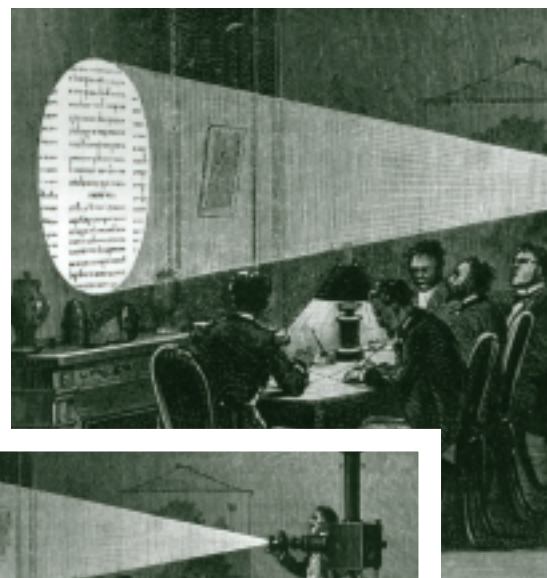
these are the very processes in which we conceptualize the most intimate of our own mental processes: remembering and forgetting. Virtually every medium for storing language has at one time been used as a metaphor for memory.

In his dialogue *Theaetetus*, Plato explains that memory is like a wax tablet, registering whatever we wish to remember. His pupils at the Academy probably carried wax tablets around, and it must have been a natural figure of speech to represent memory as a writing surface, the quality of which varied with the composition of the wax.

In the Middle Ages, as Mary Carruthers shows in *The Book of Memory* (Cambridge University Press, 1990), paper scrolls and books became the dominant metaphors for memory. After the death of Thomas Aquinas in 1274, his secretaries testified that his prodigious memory seemed to be incapable of forgetting anything, "as if knowledge were ever increasing in his soul, as page is added to page in the writing of a book".

In modern times, Freud compared human memory to a magic slate, with a surface layer that is periodically erased and therefore has unlimited capacity, and a deeper layer of wax which retains permanent traces of our experience. From the 1970s onwards, the most prominent model of memory has been the computer. When Alan Turing called the capacity of a computer to store information its "memory", he added that the machine had to "read" data from a tape.

Plato, Aquinas, Freud, Turing and present-day researchers obviously have widely diverging interpretations of memory. But what they all have in common is that they conceptualized the hidden processes of remembering and forgetting in terms of a metaphor that invokes a series of concrete, comprehensible images. This is precisely what metaphors are designed to do — they are not simply passive instruments of thought. Metaphors, particularly the powerful ones such as reading and writing, lay the tracks for our trains of associations. Whether we are investigating the mechanisms of memory or genetics, metaphors direct the way we think, name and hypothesize. Once



Microphotography found an early application in the military. During the German siege of Paris in 1871, messages were photographed on extremely thin sheets of collodion and sent off using carrier pigeons. The procedure required a specially equipped reading room for the enlargement and transcription of dispatches. Illustration from *De Natuur* (1889), at the time the Dutch equivalent of *Nature*.

we start thinking along the lines of reading and writing, we will end up with memory traces as 'engrams' and cells containing 'genomic libraries'.

Is this a cause for concern? Perhaps one should not worry about the inevitable. In 1666, when Samuel Parker, a member of the Royal Society of London, advocated a general ban on the use of figurative language in scientific discourse, he expressed a particular horror at metaphors, whose "wanton and luxuriant fancies climbing up into the Bed of Reason, do not only defile it by unchast and illegitimate Embraces, but instead of real conceptions and notices of things impregnate the mind with nothing but Ayerie and Subventaneous Phantasmes".

But since then, neither geneticists nor psychologists seem to have heeded his warning. The fact of the matter is that we simply cannot escape metaphorical thinking. It comes as naturally to us as, indeed, it came to Parker. ■

Douwe Draaisma is in the Department of Theory and History of Psychology, State University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands.



# Looking down is looking up

Jack M. Loomis

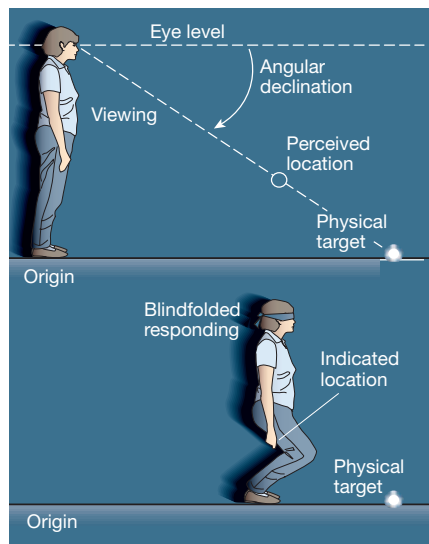
How do we perceive distance using only one eye? A neat variation on existing methods of measuring visually perceived distance highlights the importance of 'angular declination', a cue long thought to be involved.

Distance perception relies on two binocular cues and many monocular cues, several of which make up the so-called pictorial cues (linear perspective and texture gradient, for instance)<sup>1</sup>. Binocular vision has received the lion's share of research attention<sup>2</sup>, in part because of its more apparent physiological basis and its many fascinating phenomena, such as the depth seen in stereoscopic photographs and random-dot posters. But the importance of monocular vision should not be underestimated: the appearance of most natural scenes varies little as one alternates between one and two eyes, and many one-eyed people lead normal lives, including participating in sports.

On page 197 of this issue<sup>3</sup>, Ooi *et al.* describe a set of beautiful experiments that provide the best evidence yet that one particular monocular cue, angular declination, is indeed a strong determinant of perceived distance. This general topic is not only of intrinsic interest. It also helps us to understand how we move and act in three-dimensional space, and is central to developing realistic computer graphics, including virtual-reality displays.

Angular declination (also known as 'height in the field') is the angle between a visual target, typically one situated on the ground, and an observer's eye level, which generally coincides with the visible horizon (Fig. 1). As an object on the ground moves closer to an observer, its angular declination increases. Ooi and her colleagues show that manipulations of angular declination do indeed cause a change in perceived distance.

To measure perceived distance, the authors extended an established methodology known as visually directed action, which 'opens the loop' between vision and action<sup>4</sup>. In blindfolded walking, for example, an observer views a target and then proceeds blind to the estimated location of the target. Walked distance is taken as the estimate of initially perceived distance. Because of concern that such a simple response might be mediated by reasoning-like processes, and so might not provide a pure measure of perception, other variants of the method have been developed<sup>5,6</sup>. In these variants, more complex spatial behaviours are used to 'triangulate' the perceived location (such as walking blind along an indirect path to the initially perceived target), thus allowing a purer measurement of perceived distance, one aspect of perceived location. However,



**Figure 1** An eye on distance — one of the experiments carried out by Ooi *et al.*<sup>3</sup>. **Top**, an observer in darkness views a glowing target on the ground. Typically, a target more than several metres away is perceived to be closer than it is, and appears slightly elevated. After viewing the target, the observer puts on a blindfold and walks forwards to the target's perceived location. **Bottom**, the observer crouches and uses her hand to indicate the perceived location of the target. Despite the neuromuscular complexity of such a response, the indicated direction is very nearly equal to the angle of declination — this is strong evidence that the response does indeed measure the perceived location, and hence its distance.

these triangulation methods require several responses to a single target.

Ooi *et al.* have developed a new variant of visually directed action that, with just a single response to a target stimulus, allows the precise determination of the target's perceived three-dimensional location (as specified by distance and two parameters of direction). In this variant, the blindfolded observer walks out to the distance of the initially perceived, and mentally updated, location of the target (Fig. 1). On stopping, the observer then provides a precise indication of the three-dimensional location with his or her hand. Although this single response might seem an obvious way of measuring perceived location, no researcher had actually used it before, perhaps out of the belief that a response of such neuromuscular complexity cannot provide a direct measurement of perception.

A striking result obtained by the authors confirms that their method does indeed measure the perceived location of a target, and so its perceived distance. Other methods of measuring perceived distance have suggested that when a visible target more than a few metres away is viewed in an otherwise dark environment, it tends to be perceived as closer than it actually is. In one of their experiments, Ooi *et al.* used their method to measure the perceived locations of single glowing targets on the ground viewed in darkness. They found that the locations indicated by the observer for more distant targets were closer than their physical locations, as well as being elevated above the ground (Fig. 1). What is remarkable, however, is that this indicated three-dimensional location, following blindfolded walk and hand positioning, is very nearly in the same direction as the initially viewed target. Given the neuromuscular complexity of the response and the fact that the indicated location is not coincident with the physical target, the close agreement between the indicated direction and the initially viewed angular declination is strong evidence that the indicated location coincides with the perceived location.

This result means that visual direction (angular declination) is correctly perceived, a notable outcome in its own right. However, of much greater interest here is perceived distance, for which the method also provides a measure. The more distant targets viewed in darkness tend to be perceived as closer than they are. But Ooi *et al.* also provide evidence that near targets viewed in darkness, and targets as far away as 7.5 metres viewed normally within a rich visual scene, tend to be perceived quite accurately.

So far, so good: the method measures perception. But the results described so far tell us nothing about angular declination as a distance cue, because other cues were available in Ooi and colleagues' experiments even when glowing targets were viewed in darkness. There have been previous indications that angular declination acts a distance cue<sup>7</sup>, but what was needed was a demonstration that manipulating angular declination alone causes a change in perceived distance.

Ooi *et al.* provide such a demonstration, by showing that manipulations of angular declination, which have no influence on other distance cues, cause variations in perceived distance. In one experiment, using

visible targets viewed against darkness, observers wore prisms that deviated the light from the targets, causing them to have larger angular declinations (signalling closer targets) than in the initial experiments described above. The walking and pointing procedure revealed a systematic increase in the perceived visual direction as well as a corresponding reduction in perceived distance.

In the same experiment, observers were allowed to adapt to the prisms beforehand by walking around a lighted room while avoiding obstacles, after which they performed blindfolded walking and pointing to targets. Adaptation to the prisms produced shifts in perceived direction towards the horizon, with corresponding increases in perceived distance. A separate experiment, in which the adaptation procedure involved throwing beanbags at a target instead of avoiding obstacles, produced similar results. This confirms that the shifts were real shifts in perceived angular declination, not some after-effect specific to the coupling between vision and walking.

The question still remained as to the exact meaning of eye-level used by the nervous system in evaluating angular declination. Visually perceived eye level (VPEL) — the direction in space that appears level with the eyes — is determined both by the visual scene and by non-visual signals, such as those from the inner ear, which indicate the gravitationally specified horizontal plane<sup>8</sup>. In an additional experiment, Ooi *et al.* asked observers to position a visible spot viewed in darkness so that it appeared at eye level. Observers carried out this procedure under the same sets of conditions used in the other experiments: with and without prisms, and while experiencing the after-effects of prism adaptation. The shifts in VPEL matched those observed with the walking and pointing procedure, indicating that angular declination is indeed measured with respect to VPEL, at least under these conditions of darkness. Another experiment, in which observers viewed a well-illuminated scene, showed that viewing through prisms both before and after adaptation produced shifts in perceived distance in the expected directions. So it seems that VPEL might also serve as the reference direction for angular declination in well-illuminated scenes.

Apart from showing the effectiveness of angular declination as a distance cue and revealing one aspect of its processing by the visual system, this research also has great methodological significance. Until now there has been no sure and simple way of measuring the perceived three-dimensional location of a viewed target, at least one beyond arm's reach. With Ooi and colleagues' unambiguous demonstration that their refinement of visually directed action does indeed measure perceived location using just a single response, the approach can be used much more widely to measure the

effectiveness of different distance cues and their integration within the visual system. ■

Jack M. Loomis is in the Department of Psychology, University of California, Santa Barbara, California 93106, USA.

e-mail: loomis@psych.ucsb.edu

1. Sedgwick, H. A. in *Handbook of Perception and Human Performance*, Vol. 1 (eds Boff, K. R., Kaufman, L. & Thomas, J. P.) 21.1–21.57 (Wiley, Chichester, 1986).

2. Howard, I. P. & Rogers, B. J. *Binocular Vision and Stereopsis* (Oxford Univ. Press, 1995).
3. Ooi, T. L., Wu, B. & He, Z. J. *Nature* **414**, 197–200 (2001).
4. Thomson, J. A. J. *Exp. Psychol. Hum. Percept. Perform.* **9**, 427–443 (1983).
5. Fukusima, S. S. *et al.* *J. Exp. Psychol. Hum. Percept. Perform.* **23**, 86–100 (1997).
6. Philbeck, J. W. *et al.* *Percept. Psychophys.* **59**, 601–612 (1997).
7. Philbeck, J. W. & Loomis, J. M. *J. Exp. Psychol. Hum. Percept. Perform.* **23**, 72–85 (1997).
8. Matin, L. & Fox, C. R. *Vis. Res.* **29**, 315–324 (1989).

### Biophysics

## Water at the nanoscale

Mark S. P. Sansom and Philip C. Biggin

You would not expect water to enter a hydrophobic carbon nanotube. But computer simulations show that it can, and studying the process should provide clues about the behaviour of biological pores.

Water continues to surprise us. Intuitively, one would not expect water to enter a narrow hydrophobic pore, such as that formed by a carbon nanotube, because of both the tube's narrowness and its 'oily', water-repellent properties. Such chemical common sense stems from our experience of the macroscopic world, however, and may not apply at the nanometre scale.

On page 188 of this issue<sup>1</sup>, Hummer and colleagues report molecular-dynamics simulations that examine the behaviour of water within carbon nanotubes, the internal diameter of which is sufficient to accommodate a single-file column of water molecules. Using standard parameters for the strength of the weak attractive force — the van der Waals' interaction — between water molecules and the carbon atoms of the nanotube, the authors show that water molecules enter the pore, forming a column five molecules in length. But if the carbon–water interaction is made a little less favourable, by reducing the depth of the interaction's energy well (the strength of the interaction) from 10.114 kcal mol<sup>-1</sup> to 10.065 kcal mol<sup>-1</sup>, there is a dramatic change in water behaviour, and the pore remains empty of water molecules for most of the time.

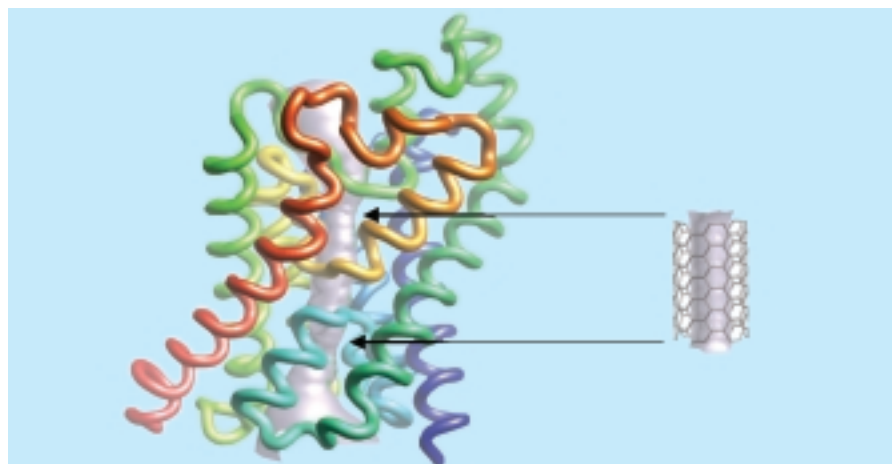
The simulations are long by current standards (50 nanoseconds or more), thereby allowing the flow of water molecules through the pore to be monitored. This work extends our understanding of how liquids behave on the nanoscale. Several studies have examined the behaviour of liquids squeezed into thin films, but comparable data on nanoscale pores are harder to come by, hence the value of simulations.

Many fluids behave abnormally when confined in a space of nanometre dimensions<sup>2</sup>. For example, simple organic liquids become solid-like when squeezed between two smooth surfaces into a film that is less than about five molecular layers thick<sup>3</sup>. In contrast, if water is squeezed between two

mica surfaces, only small changes in viscosity occur<sup>4</sup>. The nature of the surfaces between which the water is confined may also have an effect. There have also been simulations<sup>5</sup> to examine the 'nano-ice' formed by water in nanotubes of different dimensions, revealing phases of ice that are not found under bulk conditions. This line of research is relevant not only to the science of carbon nanotubes. Water-filled pores of similar dimensions to nanotubes, for instance aquaporins and ion channels, are present in many membrane-spanning transport proteins (Fig. 1). Aquaporins form water-permeable pores in many cell membranes, and ion channels form ion-permeable pores that govern the electrical properties of nerve cells.

Hummer *et al.*<sup>1</sup> describe simulations of a narrow — (6,6), in trade notation — carbon nanotube in a 'box' of 1,000 water molecules. A crucial feature of the simulations is that water molecules are allowed to enter and leave the tube freely, rather than simply being confined within it. Perhaps the most interesting aspect of the simulations is how the wet–dry transition (that is, the entry and exit of water to and from the pore) can be manipulated, and the consequences of such manipulation. In the simulations, the transition could be induced by a minute change in the strength of the water–carbon interaction potential, implying that small changes in the polarity (electrostatic state) and/or geometry of a pore might drive transitions between full and empty states.

How is water able to enter a hydrophobic channel? The key lies in the energetics of hydrogen bonding. If water molecules enter the nanotube, they each lose on average two out of their four hydrogen bonds. This would be expected to be too high an energetic cost to allow the water to enter. However, fluctuations in the number of hydrogen bonds per water molecule in the bulk aqueous phase mean that a significant fraction of water molecules are incompletely hydrogen-



**Figure 1 Pore comparisons.** Left, a biological pore (grey), as found in the bacterial water and glycerol channel, GlpF. The pore is formed by the side chains of the protein backbone (seen here as coloured spaghetti). Right, a similar pore in a narrow carbon nanotube, such as that for which Hummer *et al.*<sup>1</sup> carried out simulations of water conduction. It is hoped that nanotubes can be used in simulations to mimic and help understand the behaviour of aquaporins and ion channels.

bonded, and thus have a low binding energy. This class of water molecules is not present within the nanotube, which provides enough difference in chemical potential to drive water into it, where the water is shielded from fluctuations. Water within the nanotube moves through the tube in bursts — an average of about 17 water molecules move right through the tube each nanosecond, but peaks of about 30 molecules  $\text{ns}^{-1}$  are observed. As expected, water moves through the tube in a highly correlated, single-file manner, maintaining the chain of hydrogen bonds.

Other simulations of water in nanotubes<sup>6</sup> have also shown that the number of hydrogen bonds within such tubes is reduced relative to conditions in bulk water. Interestingly, in the narrowest (6,6) nanotube, water diffuses about two times faster along the tube than it does in the bulk. This contrasts with the situation in other, more polar pores, in which water diffusion is reduced<sup>7</sup>, and shows the importance of the exact nature of the pore–water interaction.

The internal radius of a (6,6) nanotube is 4.1 Å; that of an aquaporin is, on average, about 2 Å. The surface lining the aquaporin pore is a mixture of hydrophobic and hydrophilic regions and so might be expected to exhibit even more complex behaviour than a ‘simple’ nanotube. The observation that entry of water into nanotubes can be fine-tuned by controlling the exact strength of water–tube interactions is directly relevant to the way in which evolution has fine-tuned different members of the aquaporin family in their relative permeability to water and glycerol<sup>8</sup>.

Hummer and colleagues<sup>1</sup> controlled the wet–dry transition of the nanotube by creating a small change in the interaction potential between the water molecules and carbon. Our own simulations<sup>9</sup> of a simple model of a

membrane pore have shown that the exit and entry of water can also be controlled by small changes in pore radius — water is excluded from a pore of 4-Å radius but is freely able to enter one of 6-Å radius. The entry and exit of water could also be controlled by the presence or absence of a small dipole in an otherwise hydrophobic pore lining, such as could be provided by a polar side chain in a biological channel. Such mechanisms for controlling pore wetting may be relevant to the mechanisms for gating — opening and closing — of biological channels.

These studies also demonstrate the power of modern computer simulations to probe the properties of water in nanoscale environments, revealing behaviour that would not be expected from our understanding of the bulk properties of water. But the water models used in such simulations have their limitations. In particular, their parameters have been set to reproduce the experimental properties of bulk water, and so we need to find out whether small changes in water models affect the simulation results.

Technologically, one of the most exciting aspects of Hummer and colleagues’ study is that it suggests that, by inducing small changes in nanotube structure, it may be possible to drive the wet–dry transition and thus to gate a synthetic nanopore. This is much like using simulations to understand the properties of biological pores at atomic resolution, because simulations allow us to move beyond the somewhat static pictures yielded by conventional structural biology. From a sociological perspective, this study confirms that some of the most exciting molecular science is happening at the interface between physics and biology. We can look forward to further fruitful encounters at this meeting point. ■

Mark S. P. Sansom and Philip C. Biggin are in the Laboratory of Molecular Biophysics, Department of



#### 100 YEARS AGO

In these days the search for some characteristic of the human body which will give unequivocal evidence of the mental nature of the individual still goes on... The result of my enquiries in this direction has been to show that only two out of the seven features of the external ear which I have investigated are correlated with a mental bias towards crime or insanity, viz. a retrograde development of the helix and a persistence of the ear tip... All that can be deduced from the present investigation is that a slightly greater proportion of the people who have ear tips and retrograde helices give themselves over to crime than those in which these two features are absent. The evidence is just sufficient to justify the suspicion that a small proportion of criminals are criminals because of their physical constitution, and it is certainly the duty of every anatomist to discover how such individuals may be recognised. As yet all the criminal marks we know of can only be stated in relative terms of the class, and have, unfortunately, no application to the individual. From *Nature* 7 November 1901.

#### 50 YEARS AGO

The two main difficulties in international collaboration... are, first, the tendency of numerous governments to express different views in different places on related international problems; and secondly, the failure of all too many governments to discharge promptly and effectively their substantive obligations as participants in a co-operative international system. The first not only creates confusion but may also affect the soundness of the plans which are formulated. The second, however, could make the formulation and execution of plans a complete waste of time. It is the growing disregard for international obligations and the unilateral repudiation of negotiated agreements that is the most disturbing feature of to-day. If it is unchecked, it can halt all schemes for economic and technical assistance to backward countries. Technicians and scientific advisers cannot work confidently and effectively apart from some rule of law and guarantee of stability, nor indeed will it be possible to recruit them for such work. If they cannot be recruited, then the programme stops. In this connexion, the challenge which present events in Persia and Egypt thus presents to international institutions does not appear to have received the attention it demands. From *Nature* 10 November 1951.



Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK.  
e-mail: mark@biop.ox.ac.uk

1. Hummer, G., Rasaiah, J. C. & Noworyta, J. P. *Nature* **414**, 188–190 (2001).
2. Granick, S. *Science* **253**, 1374–1379 (1991).
3. Klein, J. & Kumacheva, E. *J. Chem. Phys.* **108**, 6996–7009 (1998).

4. Raviv, U., Laurat, P. & Klein, J. *Nature* **413**, 51–54 (2001).
5. Koga, K., Gao, G. T., Tanaka, H. & Zeng, X. C. *Nature* **412**, 802–805 (2001).
6. Marti, J. & Gordillo, M. C. *Phys. Rev. E* **64**, 021504 (2001).
7. Tieleman, D. P., Biggin, P. C., Smith, G. R. & Sansom, M. S. P. *Q. Rev. Biophys.* (in the press).
8. Fu, D. *et al. Science* **290**, 481–486 (2000).
9. Beckstein, O., Biggin, P. C. & Sansom, M. S. P. *J. Phys. Chem. B* (in the press).

## Alzheimer's disease

# An inflammatory drug prospect

Bart De Strooper and Gerhard König

There is no known cure for Alzheimer's disease. But new hope (for mice at least) comes from an in-depth investigation of a class of drugs used to treat inflammatory diseases.

Some of the drugs used to treat diseases such as rheumatoid arthritis have an unexpected benefit: they reduce the risk of Alzheimer's disease<sup>1,2</sup>. Until now it was thought that the anti-inflammatory properties of the drugs could explain this odd side effect. But on page 212 of this issue Weggen and colleagues<sup>3</sup> offer a different explanation, based on their studies of mice and cultured cells. If the findings can be extended to people, these drugs could eventually form the basis of treatments for this devastating neurodegenerative disorder.

Alzheimer's disease is characterized by several changes in the brain, including the build-up of protein deposits known as amyloid plaques outside nerve cells, and the breakdown of neurons. The plaques consist of protein fibres, some 7–10 nanometres thick, that are mixed with small peptides called amyloid- $\beta$  (Ab) peptides. Mature plaques also contain degenerating nerve endings, and are surrounded by active astrocytes and microglia — cells that help to clear up debris in the brain.

The presence of these cells and certain inflammatory proteins has suggested that inflammation contributes to Alzheimer's disease. Moreover, 'retrospective' epidemiological studies<sup>1,2</sup> have shown that patients who are taking certain nonsteroidal anti-inflammatory drugs (NSAIDs) have a lower than normal risk of developing this disorder. But, as Weggen *et al.*<sup>3</sup> show, it is not only the anti-inflammatory effects of the compounds that are important here. Instead, the authors propose that some of these drugs directly affect the generation of the aggressive 42-amino-acid-long form of the Ab peptide, Ab42.

This Ab42 peptide, together with a series of slightly shorter and longer variants, is generated by the cleavage of the amyloid precursor protein (APP; Fig. 1) — a membrane protein with unknown function that is expressed in a wide range of tissues. Cleavage is carried out by enzymes that were until recently known only by aliases:  $\beta$ -secretase and  $\gamma$ -secretase. But considerable progress

has now been made in identifying the proteins behind these names<sup>4,5</sup>. For instance, the  $\gamma$ -secretase is apparently a multiprotein complex, containing presenilin proteins, that cleaves APP in its transmembrane domain<sup>4,5</sup>. It also cleaves other transmembrane molecules, most importantly Notch<sup>6</sup> — a protein that is crucial during development.

The  $\gamma$ -secretase complex can cleave APP at different positions in its transmembrane domain, which explains why Ab peptides of different lengths are generated (Fig. 1). In patients with the inherited form of Alzheimer's disease, mutations in the APP gene or in the presenilin genes often occur, and these mutations cause small but significant changes in the preferred cleavage site in APP<sup>7</sup>. This increases the production of the Ab42 peptide relative to the other major species — the slightly shorter Ab40. This subtle change in ratio is thought to be at the root of Alzheimer's disease<sup>4,5</sup>, because the Ab42 peptide precipitates much more

quickly than Ab40. It seems plausible, then, that drugs that reduce Ab42 levels might prove useful in treating Alzheimer's disease. However, recently described inhibitors of  $\gamma$ -secretase decrease the generation of all Ab-peptides and, more problematically, the cleavage of Notch as well<sup>6,8</sup>.

The new results<sup>3</sup> hint that NSAIDs might control  $\gamma$ -secretase specifically to block the generation of Ab42 only. The authors tested the effects of three NSAIDs — ibuprofen, indomethacin and sulindac sulphide — on the processing of APP in cultured cells, and found that the ratio of Ab42 to Ab40 was substantially decreased. Moreover, ibuprofen lowered the levels of Ab42 in the brains of mice that had been engineered to provide a model of Alzheimer's disease (described in ref. 9). Importantly, these drugs do not seem to interfere with the processing of Notch<sup>3</sup>.

The results shed light on why different anti-inflammatory drugs have different effects on Alzheimer's disease<sup>10</sup>. As mentioned above, several retrospective studies suggested that some NSAIDs may be beneficial<sup>1,2</sup>. In a small 'prospective' trial, indomethacin slowed the cognitive decline seen in Alzheimer's patients<sup>1,2</sup>. And ibuprofen reduces both amyloid plaques and inflammation in mouse models of the disease<sup>9</sup>. But clinical trials using other classes of anti-inflammatory compound, such as hydroxy-chloroquine<sup>11</sup> or the steroid prednisone<sup>12</sup>, did not show clear benefits to Alzheimer's patients. Similarly, the best known NSAID — aspirin — had an unconvincing effect in epidemiological studies<sup>10</sup>. This all seemed rather puzzling at the time: if inflammation is involved in the progression of Alzheimer's disease, why would one anti-inflammatory drug benefit patients while another did not? Weggen *et al.*'s results offer an answer: we



Figure 1 Alzheimer's and amyloid. The lower part of the figure shows amino-acid sequences in single-letter code. Amyloid- $\beta$  peptides (Ab) come in a variety of sizes, of which the 42-amino-acid form (Ab42) is thought to contribute significantly to the development of Alzheimer's disease. These peptides are produced from the amyloid precursor protein (APP) by enzymatic activities known as  $\alpha$ -,  $\beta$ - and  $\gamma$ -secretases. The  $\gamma$ -secretase (or  $\gamma$ -secretases) can cleave APP at several positions, generating Ab peptides of different lengths (38, 40 or 42 amino acids). This relaxed enzymatic specificity is poorly understood. Asterisks indicate the positions of mutations in the APP gene that correlate with the development of Alzheimer's disease; most of these mutations cause the balance of Ab peptides to shift in favour of Ab42. The numbers 671, 687 and 713 indicate the positions of the amino acids in the protein chain. The p3 peptide is generated by combined  $\alpha$ - and  $\gamma$ -secretase activity.

| Therapy                                                                           | Mechanism                                                                                                                                     |
|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|
| Vaccination therapy                                                               | Stimulates immune response against A $\beta$ peptides <sup>13</sup> , leading to the peptides being cleared from the body                     |
| $\beta$ - and $\gamma$ -secretase inhibitors                                      | Inhibit $\beta$ - or $\gamma$ -secretases <sup>4</sup> , leading to decreased production of A $\beta$ peptides                                |
| Cholesterol-lowering drugs (statins)                                              | Increase processing of amyloid precursor protein by $\alpha$ -secretase <sup>14</sup> , leading to decreased production of A $\beta$ peptides |
| Copper/zinc chelator (clioquinol)                                                 | Decreases the interaction of copper and zinc with A $\beta$ peptides <sup>15</sup> , leading to amyloid plaques being cleared                 |
| Nonsteroidal anti-inflammatory drugs (ibuprofen, indomethacin, sulindac sulphide) | Selectively inhibit the production of the A $\beta$ 42 peptide <sup>3</sup>                                                                   |

Figure 2 Strategies for reducing the levels of amyloid peptides in the body. All have been validated in animal models.

need to take into account the effect of the drugs on the generation of A $\beta$ 42. It is surely no coincidence that positive clinical effects were seen mainly with those NSAIDs that are now known to block A $\beta$ 42 production. Aspirin, for instance, does not appear to affect this process<sup>3</sup>.

As usual, a word of caution is appropriate. One potential source of concern is the increase in levels of the A $\beta$ 38 peptide<sup>3</sup> that accompanies the decrease in A $\beta$ 42 — it is not known whether or not the shorter peptide is harmful. Moreover, although many investigators think that the A $\beta$  peptides are important in neurodegeneration, definitive proof of this 'amyloid hypothesis' is lacking. So it remains uncertain whether reducing the amyloid burden will be valuable in treating Alzheimer's disease. But the new work adds to the list of 'anti-amyloid-peptide' strategies (Fig. 2)<sup>4,13–15</sup>, which might be useful clinically and will also provide a definitive test for the hypothesis.

From both fundamental and practical standpoints, we now need to find out the mechanism behind Weggen *et al.*'s observations. The anti-inflammatory, pain-killing and fever-reducing activities of the NSAIDs — as well as their major detrimental side effects — result mainly from their inhibition of the enzymes cyclooxygenase-1 and -2. But Weggen *et al.* found no correlation between such inhibition and the decreased generation of A $\beta$ 42, suggesting that the A $\beta$ 42-lowering activity does not depend on cyclooxygenases. One wonders whether this activity results from NSAID-driven conformational changes in the g-secretase complex. More potent and specific analogues of these drugs are now

needed so that we can establish why certain NSAIDs modulate g-secretase activity while others do not.

But we may not have to wait for this next generation of compounds before patients can benefit. Researchers already have a wealth of clinical experience with NSAIDs, and the concentrations that have effects on A $\beta$ 42 in mice or cell cultures<sup>3</sup> are already being used to treat inflammatory diseases in the clinic. We can at least hope that researchers will soon be testing the effects of these drugs on Alzheimer's disease in clinical trials.

Bart De Strooper is at the Flanders Interuniversity Institute for Biotechnology and the K. U. Leuven, B-3000 Leuven, Belgium.

e-mail: bart.destrooper@med.kuleuven.ac.be

## Microbiology

# Tackling anthrax

Arthur M. Friedlander

Antibiotic development is the first priority in responding to terrorist use of anthrax. But structural studies offer new leads in the hunt for more effective anti-toxin treatments.

In 1990, during the Gulf War, and then again in 1998, the decision was made to vaccinate US military personnel against the possible use of anthrax as a biological weapon. Never before had vaccinations been carried out as a response to anything but the natural occurrence of a disease. Now, spores of the causative bacterium, *Bacillus anthracis*, are being used to kill civilians and create panic: the question of how to deal with anthrax has shot straight to the top of the medical agenda.

There are three ways to tackle the disease: vaccination to prevent bacterial infection in the first place; antibiotics, to attack infection if it occurs; and anti-toxin treatments for the bacterium's toxic effects. Papers by Bradley *et al.*<sup>1</sup> and Pannifer *et al.*<sup>2</sup> on pages 225 and 229 of this issue, which follow their on-line publication on 23 October, will help in developing the last approach. The possible emergence of antibiotic- and vaccine-resistant organisms<sup>3</sup> adds further impetus to the search for new approaches to preventing or treating anthrax.

The existence of toxins of *B. anthracis* was postulated by Robert Koch in the nineteenth century, when he discovered the cause of anthrax. Since their subsequent discovery almost 50 years ago, the toxins (known as 'lethal' and 'oedema'), along with the capsule that surrounds the bacterium, have been recognized as the main components that enable anthrax to cause harm. However, the manifestations of anthrax infection are not solely due to the effects of the toxins, as is the case with diphtheria, tetanus or botulism.

Rather, in anthrax the bacterium invades and grows to high concentrations in the host; the toxins act mainly by damaging defensive

Gerhard König is at Bayer AG, Pharma Research CNS, 42096 Wuppertal, Germany.

e-mail: gerhard.koenig.gk@bayer-ag.de

- McGeer, P. L., Schulzer, M. & McGeer, E. G. *Neurology* **47**, 425–432 (1996).
- Breitner, J. C. *Annu. Rev. Med.* **47**, 401–411 (1996).
- Weggen, S. *et al. Nature* **414**, 212–216 (2001).
- Dominguez, D. I., De Strooper, B. & Annaert, W. *Amyloid* **8**, 124–142 (2001).
- Selkoe, D. J. *Nature* **399**, A23–A31 (1999).
- De Strooper, B. *et al. Nature* **398**, 518–522 (1999).
- Scheuner, D. *et al. Nature Med.* **2**, 864–870 (1996).
- Li, Y. M. *et al. Nature* **405**, 689–694 (2000).
- Lim, G. P. *et al. J. Neurosci.* **20**, 5709–5714 (2000).
- Jones, R. W. *Lancet* **358**, 436–437 (2001).
- Van Gool, W. A. *et al. Lancet* **358**, 455–460 (2001).
- Aisen, P. S. *et al. Neurology* **54**, 588–593 (2000).
- Schenk, D. *et al. Nature* **400**, 173–177 (1999).
- Fassbender, K. *et al. Proc. Natl Acad. Sci. USA* **98**, 5856–5861 (2001).
- Cherny, R. A. *et al. Neuron* **30**, 665–676 (2001).

cells called phagocytes, causing the immune system to malfunction. Late in the infection, toxins may be present in large amounts in the blood and contribute directly to the death of the infected organism. Some studies, in non-human primates, suggest that lethal toxin by itself is not even particularly potent, requiring milligram quantities to cause death. But the results of other work<sup>4</sup>, in mice, imply that further non-toxin components contribute to virulence that have yet to be identified. So antibiotics constitute the mainstay of treatment, although anti-toxins have long been considered<sup>5</sup> an essential 'adjunctive' therapy, and remain so.

The toxins are composed of three proteins: a cell-receptor binding protein, known as protective antigen; and two enzymes, lethal factor and oedema factor. The papers by Bradley *et al.*<sup>1</sup> and Pannifer *et al.*<sup>2</sup> respectively report on the identity of the cellular receptor for protective antigen, and the crystal structure of lethal factor. This valuable information about the toxins will allow the identification of vulnerable targets for anti-toxin therapy.

Lethal factor is a zinc protease, a type of enzyme that contains zinc and cleaves other proteins. Oedema factor belongs to a class known as adenylate cyclases. When combined with lethal factor, protective antigen constitutes lethal toxin; with oedema factor it makes oedema toxin. From cell-culture studies it seems that anthrax operates as follows. First, protective antigen is cleaved, and so activated, by a protease on the surface of the cell under attack. It then forms heptamers — aggregations of seven — and

subsequently binds one or more molecules of lethal or oedema factor, or both. The complex passes into the cell through the receptor for protective antigen and on into an acidic compartment inside the cell. There the heptamer inserts into the compartment's membrane, releasing lethal and oedema factors into the cell body where they exert their toxic effects. The precise molecular targets remain unknown.

Several features of these events remain unclear, such as the ratios of molecules in the complex between protective antigen and lethal and oedema factors, and whether both factors are included. A protease in the blood stream can also activate protective antigen, and complexes with lethal factor occur in the blood of infected animals. So the relative importance of the two proteases in toxin action *in vivo* is unknown.

The structure of lethal factor<sup>2</sup> will help in the identification of drugs that interfere with its binding, and maybe that of oedema factor, to protective antigen; indeed, such a peptide inhibitor has been described<sup>6</sup>. Investigating inhibitors of the activities of the two factors themselves is another route, and will be aided by knowledge of the crystal structures<sup>2</sup>. Therapies might include soluble toxin receptors and other drugs to prevent protective antigen from binding to its receptor. Non-toxic mutants of protective antigen have been shown to neutralize toxin<sup>7,8</sup>, and inhibitors of the protease(s) that activates it might have the same effect. Here, detailed knowledge of toxin kinetics during infection will be required, and the timing of drug delivery is critical. Another tactic may develop from understanding how a recently discovered motor protein confers resistance to lethal toxin in some phagocytes<sup>9</sup>.

Further adjunctive treatments might include antibodies to the toxin, spore and bacillus. Other bacteria that cause sepsis exert pathological effects by triggering an inflammatory response from molecules known as cytokines; treatment with activated protein C has been reported<sup>10</sup> to be effective in such cases. The possible role of cytokines in anthrax warrants further evaluation, as do other ways of preventing the physiological consequences of the toxins. The extensive experience of testing adjunctive therapies for sepsis will be invaluable in guiding this work. In the current circumstances, however, priority should go to evaluating antibacterial drugs, initially antibiotics that have already been licensed for use.

Two final points are that information from sequencing of the *B. anthracis* genome, currently underway, may prove invaluable in tackling anthrax. And it is extraordinarily difficult to test anti-anthrax therapies in humans, so large clinical trials with non-human primates may well be needed.

In the early days of microbiology, 125 years ago, anthrax was significant mainly as an

economically damaging disease of domesticated animals. The world's scientific community addressed that problem and developed effective countermeasures. It is now necessary once again to focus on anthrax, along with other pathogenic microorganisms, this time as agents of biological terrorism and threats to civilization. During the Second World War, the Office of Scientific Research and Development in the United States, and similar agencies in other countries, coordinated the application of science to warfare<sup>11</sup>. The same level of organization and commitment is needed today.

Arthur M. Friedlander is at the US Army Medical Research Institute of Infectious Diseases, 1425

Porter Street, Fort Detrick, Frederick, Maryland 21702, USA.

e-mail: arthur.friedlander@amedd.army.mil

1. Bradley, K. A., Mogridge, J., Mourez, M., Collier, R. J. & Young, J. A. T. *Nature* **414**, 225–229 (2001).
2. Pannifer, A. D. *et al.* *Nature* **414**, 229–233 (2001).
3. Pomerantsev, A. P. *et al.* *Vaccine* **15**, 1846–1850 (1997).
4. Welkos, S. L. *Microb. Pathogen.* **10**, 183–198 (1991).
5. Lincoln, R. E. & Fish, D. C. in *Microbial Toxins* (eds Montie, T. C., Kadis, S. & Ajil, S. J.) 361–414 (Academic, New York, 1970).
6. Mourez, M. *et al.* *Nature Biotechnol.* **19**, 958–961 (2001).
7. Sellman, B. R., Mourez, M. & Collier, R. J. *Science* **292**, 695–697 (2001).
8. Singh, Y., Khanna, H., Chopra, A. P. & Mehra, V. J. *Biol. Chem.* **276**, 22090–22094 (2001).
9. Watters, J. W., Dewar, K., Lehoczy, J., Boyartchuk, V. & Dietrich, W. F. *Curr. Biol.* **11**, 1503–1511 (2001).
10. Bernard, G. R. *et al.* *N. Engl. J. Med.* **344**, 699–709 (2001).
11. Singer, M. *Washington Post* A21 (24 September, 2001).

## Papermaking

# Green chemistry through the mill

Terry Collins

Turning wood into paper uses lots of chemicals, whose waste products are a serious environmental concern. A new approach to the problem conjures up some clever chemistry but shows that there are no quick fixes.

The pulp and paper industry is a foundation stone of the developed world. It is hard to imagine the dramatic expansion in education over the past two centuries having occurred without paper. However, wood pulping takes billions of tonnes of raw materials from the ecosphere each year and turns them into waste. The enormous scale of the processes and the intimate contact they have with the atmosphere and oceans place the industry at the front line between the human economy and the ecosphere. So there are both economic and environmental reasons to run it in a sustainable manner.

Perhaps nowhere is this more important than in pulp delignification, the process by which lignin — the glue-like substance holding the wood fibres together — is removed from wood pulp. On page 191 of this issue, Ira Weinstock and colleagues<sup>1</sup> describe a new delignification catalyst that the authors hope will be more environmentally friendly than existing technologies. The process they have developed is chemically interesting, but many problems still remain.

Raw wood pulp consists primarily of three polymers: gluey lignin, and fibre-forming cellulose and hemicellulose. The key to producing high-quality white paper is the selective removal of lignin, which causes pulp to be dark brown<sup>2</sup>. Around 100 million tonnes of this type of pulp is delignified and bleached annually. The goal of chemical delignification is to use an oxidizing agent to selectively degrade lignin without attacking the other solid cellulose. This is not easy, however, because there are few inexpensive oxidants that can selectively attack lignin. During oxidation, electrons are removed from lignin, causing it to

break apart and become water soluble and easy to separate from the cellulose by washing. Throughout most of the past century, elemental chlorine (Cl<sub>2</sub>) was the principal delignifying and bleaching agent used worldwide. But this technology generates chlorinated waste products, such as polychlorinated dibenzodioxins and -furans (PCDD/Fs). Over the past few decades it has become clear that some PCDD/Fs have endocrine-disrupting effects<sup>3,4</sup> and other severe toxic properties, forcing the industry to move away from using chlorine.

Pulp bleaching in developed countries is now largely based on chlorine dioxide (ClO<sub>2</sub>), which produces fewer chlorinated pollutants. One can only salute the pulp and paper industry for making the huge investment to retool, but ClO<sub>2</sub> almost certainly does not represent the final word in improving delignification. The capital and operating costs of ClO<sub>2</sub> pulp mills are considerable. Moreover, ClO<sub>2</sub> production is energy intensive, undermining its long-term economic attractiveness. These factors conspire to place ClO<sub>2</sub> out of reach as an alternative to chlorine for much of the developing world. Yet toxic PCDD/Fs are volatile, persistent and accumulate in the global ecosphere and food chain, irrespective of where they are produced<sup>5</sup>.

From this perspective, one can see that new delignifying technologies must have low capital and operating costs, including low energy requirements, and should exhibit comparable or higher selectivity than ClO<sub>2</sub> throughout the entire bleaching process, while eliminating environmental concerns. This presents a huge chemical and technological challenge. One approach is to develop catalytic agents to improve the oxidizing



power of chlorine-free oxidants, such as oxygen, hydrogen peroxide or ozone, at low temperatures. In environmental and economic terms, oxygen is the most desirable of these, followed by hydrogen peroxide. These are the main oxidants used in nature, and react with wood pulp to yield natural waste products. Oxygen is effective in the early stages of delignification when there is plenty of lignin in the pulp, but when lignin concentration is low the high reactivity of oxygen leads to indiscriminate degradation of the cellulose fibres. Other chlorine-free bleaching technologies have also been developed, but these have found only partial favour in the industry because of high cost and low selectivity.

This is where Weinstock *et al.*<sup>1</sup> come in. They have developed a two-step delignification process based on a catalyst for use with oxygen in water. The catalyst is a polyoxometallate cluster ion (POM),  $[\text{AlVW}_{11}\text{O}_{40}]^{61}$ , which in the first step of the process extracts an electron from lignin in unbleached pulp. This step requires heating to 130 °C for three hours, and high selectivity for lignin is claimed. Because lignin oxidation produces protons, the authors have to control the pH of the solution, which they do by adding other POMs to the mixture. The way the various POMs shuffle their identities while keeping the pH neutral is truly ingenious. In the second step, the POM mixture created in the first step is reoxidized by oxygen and returned to its original composition; the soluble lignin fragments are simultaneously oxidized, turning them into  $\text{CO}_2$  and  $\text{H}_2\text{O}$ . Together these two steps result in the transfer of electrons from lignin to oxygen. The delignified pulp must then be separated from the reaction mixture and washed free of POMs and soluble lignin fragments.

Does this clever chemistry presage a desirable new technology? I have my doubts. The biggest chemical drawback is that the POM is not behaving as a catalyst in the usual sense — a good catalyst would be able to rapidly extract thousands of electrons from lignin, whereas each POM is capable of extracting only one electron per reaction cycle. And because only 15% of the active POM is used, the actual number is 0.15 electrons per POM ion per reaction cycle. By comparison,  $\text{ClO}_2$  extracts five electrons from lignin and usually 100% of the reagent is active.

Apart from their low efficiency, the POMs are massive, some 40 times heavier than  $\text{ClO}_2$ , so enormous quantities would be needed. In fact, the POM:pulp ratio in the delignifying mixture described by Weinstock *et al.* is about 170:1 by weight. Such an oxidation can be likened to one person commuting to work in a diesel locomotive. The ratio will probably improve as the authors search for more efficient POMs, but an industrial mill working to transform hundreds of tonnes of pulp per day will need thousands of tonnes of POM.

More worrying still, from an environ-

mental perspective, are the heavy-metal components of POMs, such as tungsten and molybdenum. The potential risks associated with using POMs on massive scales are unclear, but molybdenosis in cattle and other mammals is a well-documented consequence of elevated levels of molybdenum in the environment. Weinstock *et al.* have designed their delignification process so that the POMs can be recovered and used again. They also claim they can optimize the efficiency of their catalyst to reduce the quantities involved. But accidental spills are unavoidable in a working pulp mill and no process is 100% recyclable, so it seems that heavy-metal contamination of mill environments and process materials would be inevitable.

Finally, from a commercial viewpoint, the POMs are expensive and would require periodic purification. As the authors acknowledge, the reaction temperatures and times are currently too high, requiring huge pressurized reactors in an industrial setting. If these problems cannot be fixed, the technology will be expensive in terms of capital and operating costs, as well as energy intensive.

By developing a catalytic process with chlorine-free reagents, Weinstock and colleagues are thinking along the right lines. But in my view there are formidable practical obstacles that diminish its commercial appeal, and serious environmental issues that need to be addressed. These problems highlight the difficulty of designing new chemical processes that meet industrial needs while ensuring a safer environment. If there is a lesson to be learned from this, it is surely that moving the chemical composition of technology further towards nature, and not away from it as in this case, is a vital strategy for green chemistry<sup>6</sup>.

Terry Collins is in the Department of Chemistry, Carnegie Mellon University, 4400 Fifth Avenue, Pittsburgh, Pennsylvania 15213, USA.  
e-mail: tc1u@andrew.cmu.edu

1. Weinstock, I. A. *et al.* *Nature* **414**, 191–195 (2001).
2. Dence, C. W. & Reeve, D. W. (eds) *Pulp Bleaching: Principles and Practice* (TAPPI Press, Atlanta, 1996).
3. Naz, R. K. (ed.) *Endocrine Disruptors: Effects on Male and Female Reproductive Systems* (CRC Press, Boca Raton, 1999).
4. Schecter, A. (ed.) *Dioxins and Health* (Plenum, New York, 1994).
5. Thornton, J. *Pandora's Poison: Chlorine, Health, and a New Environmental Strategy* (MIT Press, Cambridge, MA, 2000).
6. Collins, T. J. *Science* **291**, 48–49 (2001).

#### Developmental biology

## Gridlock in the blood

Gavin Thurston and George D. Yancopoulos

All blood vessels originate from the same precursor cells in early embryos. So how do those precursors decide whether to contribute to arteries or veins? Studies of zebrafish bring us closer to the answer.

How do developing embryos grow a complex network of blood vessels with the correct proportions of arteries, veins and capillaries? On page 216 of this issue Zhong and colleagues<sup>1</sup> describe the crucial role of a gene called *gridlock*, which influences the earliest decisions that the precursors of blood vessels must make.

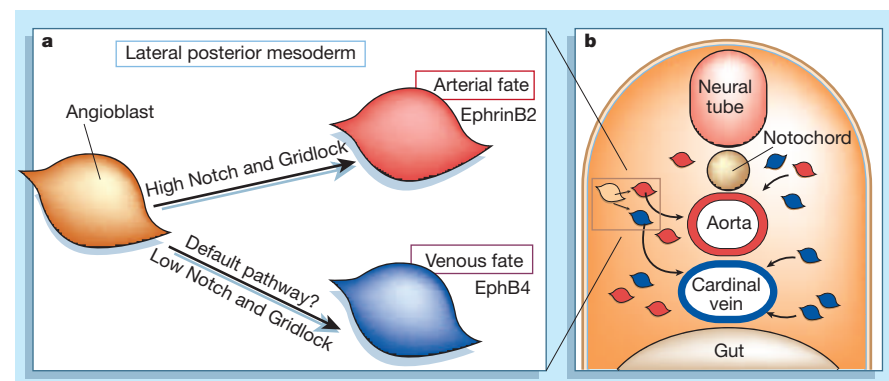
The first vessels formed in a vertebrate embryo are derived from precursor cells called angioblasts, which are initially found at the periphery of the embryo in tissue known as the lateral posterior mesoderm (LPM). The angioblasts migrate to the midline — the longitudinal axis that marks the centre of the embryo — and coalesce to form the primitive aorta and early veins. It was once thought that many of the differences between arteries and veins arise after the onset of blood flow, in response to differences in pressure and flow. But it is now clear that primitive vessels carry information about their identity before flow begins. For example, vessels that are fated to become arteries express ephrinB2, a member of a large class of transmembrane proteins<sup>2</sup>. Cells with a venous fate, by contrast, express Eph4, the receptor for ephrinB2. So how is an arterial versus a venous fate determined?

Zebrafish have become invaluable to those

who study early embryonic development. Their embryos are transparent, which makes it easy for researchers to visually screen randomly generated mutants for interesting anatomical defects. It is also becoming easier to clone the genes mutated in these defective embryos. Further analyses are aided by molecular tools known as modified antisense oligonucleotides, which can be used to reduce the levels of a protein of interest; such tools have not been adequately developed for use in higher organisms. Moreover, zebrafish have particular advantages for studies of cardiovascular defects — the embryos are unusual in that they can survive without blood flow.

Several years ago, a screen for zebrafish with defects in their cardiovascular systems led researchers to a mutant that they dubbed 'gridlock', because its major blood vessels failed to form<sup>3</sup>. The normal *gridlock* gene is expressed first in the LPM region in which angioblasts are located, and later in the developing aorta — but not in developing veins<sup>4</sup>. The Gridlock protein is related to a class of gene-expression (transcriptional) repressors that seem to have key roles in cell-fate decisions<sup>4</sup>. Zhong *et al.*<sup>1</sup> investigated whether Gridlock, too, has a say in cell fate.

The authors<sup>1</sup> started by labelling angio-



**Figure 1** How blood vessels develop. **a**, Arteries and veins are formed from precursor cells called angioblasts, which are initially found in the lateral posterior mesoderm (a tissue in the periphery of the embryo). Zhong *et al.*<sup>1</sup> show that, in zebrafish, angioblasts that express the proteins Notch and Gridlock will contribute to arteries. Those that do not express these proteins will contribute to veins. **b**, Once committed to forming arteries or veins, the cells migrate towards the midline and coalesce to form the appropriate blood vessel. The figure shows a cross-section of an embryo.

blasts in the LPM with a fluorescent tag and monitoring their fate. This appeared to be predetermined: all the progeny of a given angioblast ended up in either the aorta or a vein, but not both. Zhong *et al.* then looked at whether *gridlock* prompts angioblasts to form arteries. They decreased levels of the Gridlock protein by using modified antisense oligonucleotides, and found a dose-dependent reduction in the size of the aorta, and an increase in the adjacent vein. Reciprocally, increasing Gridlock activity by injecting *gridlock* messenger RNA into early embryos decreased the size of the vein (though it did not increase the size of the aorta). The implication is that angioblasts make up a limited pool of precursors whose default fate is venous, and that the expression of *gridlock* suppresses the venous characteristics and helps to drive the precursors towards forming an aorta.

Transcriptional repressors that are related to Gridlock can be expressed in response to signals from the Notch family of receptor proteins. To find out whether this applies to Gridlock as well, Zhong *et al.* injected either an activated form of Notch1 or an inhibitor of the Notch1 signalling pathway into zebrafish embryos. They found that activation of the pathway by Notch1 induced the expression of *gridlock*, whereas suppression of the pathway both reduced *gridlock* expression and disrupted the formation of the aorta.

Another study<sup>5</sup>, by Lawson *et al.*, had already implicated the Notch pathway in arterial fate, particularly in the expression of artery-specific proteins such as ephrinB2. But these authors found that certain arterial characteristics — such as the expression of ephrinB2 — could be ablated even when *gridlock* expression was maintained, raising questions about where *gridlock* comes into artery determination. These findings can be reconciled with those of Zhong *et al.* if one assumes that the response to *gridlock* is dose dependent, as indeed Zhong *et al.*'s results suggest. So, the total absence of *gridlock*

results in the default venous fate and obliteration of the aorta. Partial expression of *gridlock* (as may have occurred in Lawson *et al.*'s study) allows some characteristics of an artery to develop, but not all, as reflected by the lack of expression of proteins such as ephrinB2.

It seems, then, that the Notch/Gridlock pathway predetermines the fate of angioblasts (Fig. 1). Of course, questions remain. How, for example, do these precursors migrate to the appropriate site and then coalesce only with similarly fated cells? The ephrin proteins and their receptors are known to be involved in the migration and sorting of other cell types, and seem likely to help maintain the identities of different blood-vessel precursors. But do they act in the LPM, in the migration pathway, or in blood vessels themselves?

Moreover, although the nature of vascular cells may indeed be predetermined by a genetic programme before blood flow begins, many studies of vessel transplantation in embryos and adults have indicated that vessels can remodel and switch their identity. Indeed, surgeons commonly transplant the saphenous vein into segments of diseased arteries. It will be fascinating to look at the involvement of proteins such as Notch1 and Gridlock during cellular reprogramming in the course of arterial disease and in transplanted veins. With this in mind, we now need to find out if the zebrafish work can be translated to mammals. Stay tuned for further gridlock alerts!

Gavin Thurston and George D. Yancopoulos are at Regeneron Pharmaceuticals, Inc., 777 Old Saw Mill River Road, Tarrytown, New York 10591, USA. e-mails: gavin.thurston@regeneron.com george@regeneron.com

1. Zhong, T. P., Childs, S., Leu, J. P. & Fishman, M. C. *Nature* **414**, 216–220 (2001).
2. Wang, H. U., Chen, Z. F. & Anderson, D. J. *Cell* **93**, 741–753 (1998).
3. Weinstein, B. M., Stemple, D. L., Driever, W. & Fishman, M. C. *Nature Med.* **1**, 1143–1147 (1995).
4. Zhong, T. P., Rosenberg, M., Mohideen, M. A., Weinstein, B. & Fishman, M. C. *Science* **287**, 1820–1824 (2000).
5. Lawson, N. D. *et al.* *Development* **128**, 3675–3683 (2001).

## Daedalus

### Vacuum in miniature

Meredith Thring once proposed a domestic robot to operate other machines, such as the vacuum cleaner, but was unable to give it human intelligence. Annoyed by the vacuum cleaner, Daedalus is now scaling it down. His ideal is a dedicated, chip-operated cleaner maybe only 3 cm across. Unlike bigger machines, it will need no operator, and will run continuously, like a diminutive pet. On tiny wheels it will nose along the carpet, sucking vigorously at everything that seems foreign to its environment. It will pack loose dust, skin particles and fibre into a little sack.

Like W. Grey Walter's old *Machina habilis*, the new microcleaner will be alert to the state of charge of its batteries. When they are low, or when its sack is full, it will head back to an illuminated or otherwise designated charging base. There the microcleaner will acquire electricity and dump its haul of rubbish, ideally in a box that can be disposed of when full.

A small vacuum device can be far more effective than a larger version, and can study possible waste more carefully. Modern chips allow a microcleaner to be quite flexible in its approach. An ultrasonic sensor will tell it where it is at all times. It will map an entire room into parallel tranches — several hundred tracks and many hundred of metres to be covered at maybe a metre every few minutes, perhaps at night if it is quiet enough. It will avoid cats, papers, feet and so on, remembering to revisit the site later. It will soon learn the regular occupants of the room (such as the furniture) and their degree of permanence.

The microcleaner should effortlessly displace bigger and clumsier machines. It will run all the time, cleaning an entire room in about a week, far more efficiently than any hand-held device. Then it will start again. Its owner will merely have to dispose of the dust and fluff it piles up at its charging-station. Unlike traditional machines it will need no operator, and it will tirelessly revisit sites temporarily denied to it.

Each room or pair of rooms will have its own cleaner. A household kitchen or dining-room will have a model that specializes in various sorts of spilt food; a factory will have one that knows how best to remove the particular swarf or waste it will find. The discriminating microcleaner will remove green fluff from a red carpet, but not from a mixed one. But mud, loose fluff, superfluous food, loose tea-leaves and metal turnings will be removed wherever they are.

David Jones

# Sheep don't forget a face

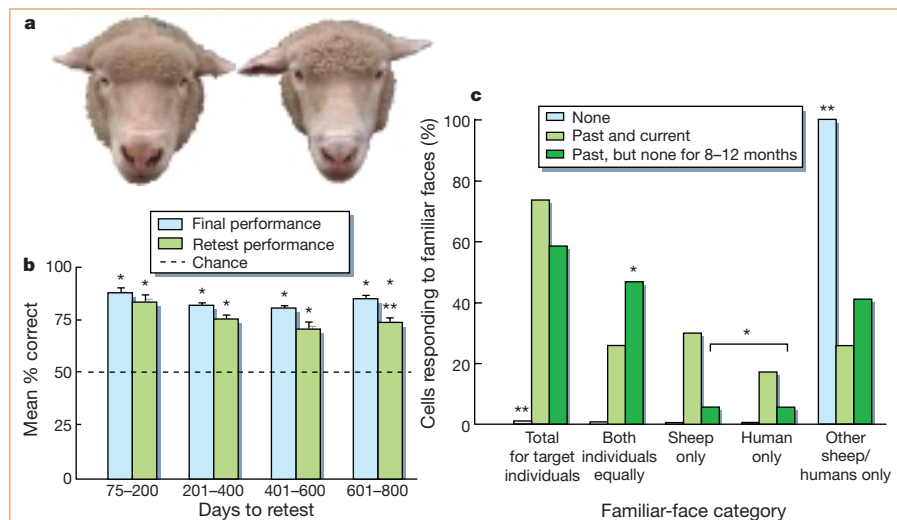
The discovery of a remarkable memory shows that sheep are not so stupid after all.

The human brain has evolved specialized neural mechanisms for visual recognition of faces, which afford us a remarkable ability to discriminate between, remember and think about many hundreds of different individuals. Sheep also recognize<sup>1–4</sup> and are attracted to<sup>1,5</sup> individual sheep and humans by their faces, as they possess similar specialized neural systems in the temporal and frontal lobes for assisting in this important social task<sup>6,7</sup>, including a greater involvement of the right brain hemisphere<sup>3,7</sup>. Here we show that individual sheep can remember 50 other different sheep faces for over 2 years, and that the specialized neural circuits involved maintain selective encoding of individual sheep and human faces even after long periods of separation.

We trained 20 sheep (*Ovis aries*) to discriminate in a choice maze between pictures showing frontal views of 25 pairs of sheep faces (Fig. 1a) by associating one member of each face pair with a food reward. Animals required 30 or more trials to reach a learning criterion of over 80% correct choice, and were given a further 400–500 trials over a period of 4–6 weeks. These latter trials confirmed that they could also discriminate between profile views of the same individuals without having to relearn the task. We then tested the sheep for retention of discrimination performance after delays of up to 800 days. Only 10 trials were given in retesting, as no face pair was ever significantly discriminated better than chance for the first time in this number of trials.

Sheep could still discriminate accurately between the 25 face pairs at all retention time points. Only after 601–800 days was performance significantly poorer than final original levels (Fig. 1b). Accurate discriminations were not restricted to original face pairings, but also occurred with three new pairs of rewarded and unrewarded faces (83.3% correct;  $P=0.0127$ , Fisher's test).

Like monkeys<sup>8</sup> and humans<sup>9</sup>, a small population of cells in the temporal and medial prefrontal cortices of the sheep brain encodes faces, as opposed to other visual objects<sup>6</sup>. A higher-order subgroup of these cells preferentially encodes the faces of one or more specific familiar individuals in the animal's current social environment. We obtained recordings from these latter cells to investigate whether they would still respond to faces of familiar individuals after a long absence. To do this, we recorded the response profiles of these cells to the faces of a specific human ('stock man') and sheep that animals were currently exposed to and



**Figure 1** Behavioural and electrophysiological evidence of long-term face recognition by sheep. **a**, Typical face pair used in behavioural tests. **b**, Mean  $\pm$  s.e.m. discriminatory performance of 20 animals on 25 different face pairs at the end of an original set of training and test trials (400–500 trials) and for 10 trials given at the indicated times after (6–7 pairs at each time point). In all cases, performance is significantly better than levels seen in the first 10 trials received for a new face pairing (mean  $\pm$  s.e.m. 451.85  $\pm$  2.4% correct; asterisk denotes  $P < 0.001$ ). Upon retesting, significant impairment compared with the end of the original set of trials was only found after a delay of 601–800 days (two asterisks denote  $P < 0.01$ , paired  $t$ -test). **c**, Proportions of cells in the right temporal and medial frontal cortices responding to different familiar faces ( $P < 0.01$  against unfamiliar faces). Histograms show the distribution of these cells' responses to the faces of two individuals (a sheep and a human) of whom animals had no experience, or 12–15 months' previous experience and were still interacting with, or 12–15 months' previous experience but had not seen for 8–12 months. In the two groups with experience of the individuals, similar proportions of cells preferentially encoded their faces, whereas the group with no experience showed no preferential encoding and had a correspondingly higher proportion of cells that responded to other familiar faces. In the group in which these two individuals had been present then absent, there was a significant reduction in the proportion of cells that responded significantly more to one or other of these faces individually ( $P < 0.05$  against all other faces) and a corresponding increase in the proportion that responded equivalently to them and to some other familiar faces (asterisk denotes  $P < 0.05$  against the group with past and current exposure; two asterisks denote  $P < 0.05$  against both groups that were familiar with the two individuals; Fisher's test).

had been previously exposed to for 12–15 months, or that they had not seen or interacted with for 8–12 months after the same length of initial exposure, or that they had never seen or interacted with. We also exposed all groups to 3–5 other individual sheep or humans for the same period.

The overall proportion of cells that responded differentially to faces, as opposed to non-face stimuli ( $P < 0.01$ ), was similar in the three groups (81 of 785,  $n=46$ ; 47 of 459,  $n=44$ ; 39 of 401,  $n=43$ ). The proportion of cells in the subgroup that responds preferentially only to the faces of all or a subset of familiar individuals ( $P < 0.01$  against 8–12 unfamiliar faces) was also similar across the groups (23 of 81; 17 of 47; 14 of 39). The proportions of these latter cells that responded to the two familiar test faces did not differ significantly in groups in which the individuals concerned were still present or had been absent for 8–12 months (Fig. 1c). However, neither of these faces was encoded preferentially by this cell type in animals who did not know the two

individuals concerned (Fig. 1c).

Sheep showed clear behavioural signs of recognizing both absent individuals by vocalizing in response to their face pictures in the same way as they did to faces of other members of their current social group. However, fewer cells responded preferentially to each face alone and more responded equivalently to both faces, as well as to those of other familiar sheep (Fig. 1c). As overall patterns of response profiles were similar in the frontal and temporal cortices, we combined the data.

Face-recognition memory may thus weaken slowly but progressively over time, owing to a shift from selective encoding of a specific individual at the highest neural level, to a level at which the individual is generalized to a category of 'familiar individual'. Presumably, failure to recognize an individual's face will therefore only occur when the proportion of the neural network that is tuned to respond preferentially to the individual face falls below a critical threshold.

The high-level encoding strategy used by



the specialized face-processing system in the sheep brain offers advantages for long-term recognition of many individuals that are similar to those for humans. In humans, analogous brain regions and neural circuits are activated equivalently when we see or form mental images of the faces of specific individuals<sup>9</sup>. This suggests that sheep may be capable of using the same system to remember and respond emotionally to individuals in their absence.

**Keith M. Kendrick, Ana P. da Costa, Andrea E. Leigh, Michael R. Hinton, Jon W. Peirce**  
Laboratory of Cognitive and Developmental Neuroscience, Babraham Institute,

Cambridge CB2 4AT, UK  
e-mail: keith.kendrick@bbsrc.ac.uk

1. Kendrick, K. M., Atkins, K., Hinton, M. R., Heavens, P. & Keverne, E. B. *Anim. Behav.* **49**, 1665–1676 (1995).
2. Kendrick, K. M., Atkins, K., Hinton, M. R., Heavens, P. & Keverne, E. B. *Behav. Proc.* **38**, 19–35 (1996).
3. Peirce, J. W., Leigh, A. E. & Kendrick, K. M. *Neuropsychologia* **38**, 475–483 (2000).
4. Peirce, J. W., Leigh, A. E., da Costa, A. P. C. & Kendrick, K. M. *Behav. Proc.* **55**, 13–26 (2001).
5. Kendrick, K. M., Hinton, M. R., Atkins, K., Haupt, M. A. & Skinner, J. D. *Nature* **395**, 229–230 (1998).
6. Kendrick, K. M. & Baldwin, B. A. *Science* **236**, 448–450 (1987).
7. Broad, K. D., Mimmack, M. L. & Kendrick, K. M. *Eur. J. Neurosci.* **12**, 731–741 (2000).
8. Perrett, D. I., Hietanen, J. K., Oram, M. W. & Benson, P. J. *Phil. Trans. R. Soc. Lond. B* **335**, 23–30 (1992).
9. Kreiman, G., Koch, C. & Fried, I. *Nature* **408**, 357–361 (2000).

## Eutrophication

# Nitrate flux in the Mississippi River

Increased delivery of biologically available nitrogen to estuaries and coastal oceans in recent decades has been linked to eutrophication and seasonal hypoxia in the northern Gulf of Mexico<sup>1,2</sup> and elsewhere<sup>3,4</sup>. We have developed a model that accounts for 95% of annual variation in delivery of nitrate to the Gulf of Mexico by the Mississippi River in 1960–98. Retrospective analysis indicates that this nitrate flux could have been reduced by 33% if the use of nitrogen-containing fertilizer in the Mississippi River basin had been cut by 12%.

Across a wide range of temperate river systems, riverine nitrogen flux has been correlated with net anthropogenic nitrogen input (NANI) to the drainage basin. NANI is defined as nitrogen input from fertilizer and by fixation and atmospheric deposition of oxidized nitrogen, minus nitrogen exported in food and feed<sup>5,6</sup>. For each unit

of NANI to a drainage basin, an average of about 0.25 units are transported to coastal waters, with the other 0.75 units being converted to gaseous compounds or stored in soil or groundwater.

Relatively little attention has been paid to temporal variation in the relationship between NANI and riverine nitrogen flux for particular river basins, which is essential for developing effective protection strategies for estuary and coastal marine ecosystems. We have focused on the temporal variation of nitrate flux in the lower Mississippi River in 1960–98, during which period a 2.5-fold increase in nitrate concentration accounted for almost all of the increase in total nitrogen concentration in the river<sup>7</sup>. A positive trend in precipitation during this period also produced a 30% increase in water yield (or stream flow), which probably enhanced nitrate delivery.

In 1960–98, NANI to the Mississippi River basin increased by roughly 80%. In the 1960s, riverine nitrate flux was 8% of NANI; by the 1990s, this figure had increased to 18%. The trend in the ratio of riverine nitrate flux to NANI is statistically significant and the ratio is significantly correlated with discharge and NANI ( $P < 0.001$ ).

By combining and adapting two earlier models that related terrestrial nitrogen input to riverine nitrogen flux<sup>5,8</sup>, we developed the following model, which accounts for 95% ( $P < 0.001$ ) of the variation in annual nitrate flux in the lower Mississippi River, including the Old River outflow, in 1960–98:

$$N_{LM} = 0.66 \times W^{0.93} \times e^{(0.13 \times \text{NANI2-5} + 0.06 \times \text{NANI6-9})}$$

where  $N_{LM}$  is the annual nitrate flux (in  $\text{kg N ha}^{-1} \text{yr}^{-1}$ ), NANI2–5 is the average annual net anthropogenic nitrogen input during the previous 2–5 years (in  $\text{kg N ha}^{-1} \text{yr}^{-1}$ ), NANI6–9 is the average annual net anthropogenic N input during the previous 6–9 years (in  $\text{kg N ha}^{-1} \text{yr}^{-1}$ ) and  $W$  is the

annual water yield (in  $\text{m yr}^{-1}$ ).

The root-mean-square error was 12% of average riverine nitrate flux and the serial correlation of residuals was not statistically significant ( $P < 0.05$ ) for lags 1–12. The results of Monte Carlo simulations for the 1980–98 period suggested that 95% of the estimation uncertainty was due to fitting the regression coefficients and 5% was due to uncertainty in components of NANI (for details, see supplementary information). The 95% confidence interval for the mean estimated nitrate flux for an individual year was  $\pm 15\%$ .

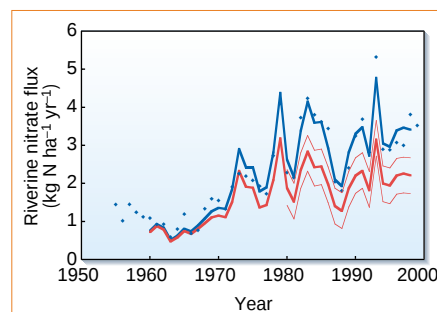
Our model suggests that changes in NANI to the Mississippi River basin influence riverine nitrate flux for the succeeding 2–9 years, although the greatest impact tends to be during the first 2–5 years. Furthermore, the observed exponential relationship suggests that small changes in NANI may lead to relatively large changes in riverine nitrate flux. This pattern may be a consequence of NANI exceeding the capacity of terrestrial and/or aquatic systems to assimilate nitrogen input.

Calculations using our equation, and made on the basis of the observed annual water yields in 1960–98, suggest that a 14.2% reduction in NANI would have led to a 33% reduction in annual average riverine nitrate flux at St Francisville in 1980–98 (Fig. 1). This reduction in fertilizer use would have been most effective in years with the greatest water yield (and therefore the greatest nitrate flux). In years with low water yield, however, there was no statistical difference between estimates of mean nitrate flux with or without a 14.2% reduction in NANI.

The 14.2% reduction in NANI could have been accomplished by a 12% reduction in nitrogen input from fertilizer if crop yields remained constant. A small reduction in such nitrogen input is unlikely to have reduced crop yields significantly, if at all. Crop yields tend to approach an economic optimum in an asymptotic manner as rates of fertilizer application increase<sup>9,10</sup>. Moreover, data concerning usage of nitrogen fertilizer suggest that there is a tendency among farmers to apply more than is necessary to achieve economically optimal production<sup>11–13</sup>.

Our findings indicate that achieving conservation goals for the Gulf of Mexico may require less reduction in fertilizer use than has been estimated from simulation modelling of edge-of-field nitrogen losses and assumptions of constant in-stream denitrification loss<sup>14</sup>. By using the available data concerning riverine nitrate transport, our analysis incorporates the effects of variation in in-stream denitrification.

The relationship between riverine nitrate flux in the Mississippi basin and NANI may continue to change as a result of



**Figure 1** Observed nitrate flux in the lower Mississippi River, including the Old River outflow, in 1955–98 (diamonds) and nitrate flux estimated using our equation (see text; black line). The thick red line shows the estimated average nitrate flux to the Gulf of Mexico if nitrogen input from fertilizer were to be reduced by 12%, assuming no reduction in crop yields. Thin red lines define the 95% confidence interval for the mean of each annual estimate. Flux values before 1960 were not included in the regression analysis because data concerning crop and livestock production data before 1950, which would be needed to calculate NANI6–9, are not readily available.

modifications in the river, in climate and in nitrogen-management practice. Continued monitoring of riverine nitrogen and NANI will refine our understanding of nitrogen dynamics in river basins and will facilitate adaptive management of conservation policies and programmes.

**Gregory F. McIsaac\*, Mark B. David\*, George Z. Gertner\*, Donald A. Goolsby†**

\*Department of Natural Resources and Environmental Sciences, University of Illinois, Turner Hall, 1102 South Goodwin Avenue, Urbana, Illinois 61801, USA

e-mail: gmcisaac@uiuc.edu

†US Geological Survey, Denver Federal Center, Denver, Colorado 80225, USA

1. Turner, R. E. & Rabalais, N. N. *Nature* **368**, 619–621 (1994).
2. Rabalais, N. N., Turner, R. E., Justic, D., Dortch, Q. & Wiseman, W. J. *NOAA Coastal Ocean Prog. Decision Anal. Ser.* **15** (Silver Spring, Maryland, 1999).
3. National Research Council 'Clean Coastal Waters: Understanding and Reducing the Effects of Nutrient Pollution' (Nat'l Acad. Press, Washington DC, 2000).
4. Diaz, R. J. *J. Environ. Qual.* **30**, 275–281 (2001).
5. Howarth, R. W. et al. *Biogeochemistry* **35**, 75–139 (1996).
6. Jordan, T. E. & Weller, D. W. *BioScience* **46**, 655–664 (1996).
7. Goolsby, D. A. & Battaglin, W. A. *Hydrol. Proc.* **15**, 1209–1226 (2001).
8. Caraco, N. F. & Cole, J. J. *Ambio* **28**, 167–170 (1999).
9. Cerrato, M. E. & Blackmer, D. G. *Agron. J.* **82**, 138–143 (1990).
10. Bullock, D. G. & Bullock, D. S. *Agron. J.* **86**, 191–195 (1994).
11. Legg, T. D., Fletcher, J. J. & Easter, K. W. *J. Prod. Agric.* **2**, 110–116 (1989).
12. US Department of Agriculture National Agricultural Statistic Service. Agricultural Chemical Usage (PCU-BB) <http://usda.mannlib.cornell.edu/reports/nassr/other/pcu-bb>
13. Taylor, H. H. *Stat. Bull.* **893** (Econ. Res. Service, US Dept Agric., Washington DC, 1994).
14. Doering, O. C. et al. *NOAA Coastal Ocean Prog. Decision Anal. Ser.* **20** (Silver Spring, Maryland, 1999).

**Supplementary information** accompanies this communication on Nature's website ([www.nature.com](http://www.nature.com)).

## Neural-network models

# Predicting spontaneous recovery of memory

Long after a new language has been learned and forgotten, relearning a few words seems to trigger the recall of other words. Neural-network models<sup>1–3</sup> indicate that this form of spontaneous recovery may result from the storage of distributed representations, which are thought to mediate human memory. Here we use a psychomotor learning task to show that a corresponding effect of spontaneous memory recovery occurs in human subjects.

Spontaneous recovery is a generic characteristic of systems in which associations are distributed over many processing units (neurons, for example)<sup>1–3</sup>. In neural-network models, after learning a set of associations, forgetting can be induced by adding noise to connections between 'neurons'. As every association depends on all connections, relearning a subset of these associations forces all connections towards their original values, resulting in

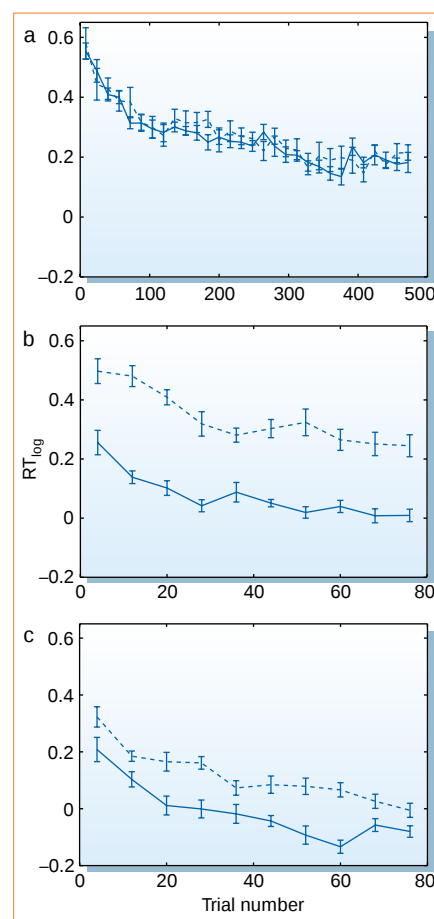
improved performance on non-relearned associations; this form of spontaneous recovery is also known as the transfer effect<sup>1,2</sup>.

The task we used to test for a transfer effect in humans involved learning to type on a keyboard on which letters had been rearranged. In each of three sessions, subjects ( $n=12$ ) were presented with an upper-case letter on each trial, and were required to press the corresponding keyboard letter. Letters were presented in random order, with an inter-trial interval of 1 second. Twenty-four letters were divided into three disjoint subsets (such as A = {ENLHUBWK}, B = {TORCFPYJ}, C = {AISDMGVX}). Subjects learned two intermixed subsets (A and B) for 480 trials (session 1). After 48 hours, subset A was relearned for 80 trials (session 2). Immediately after session 2, subjects were tested for 80 trials on subset B (session 3). The protocol in all three sessions was identical.

We predicted that, after initially learning subsets A and B (session 1), relearning subset A (session 2) would facilitate performance on the non-relearned subset B (session 3). Accordingly, we compared reaction times for subset B (session 3) in this transfer condition with those in a control condition. In the control condition, instead of relearning subset A in session 2, subjects learned a 'new' subset, C.

Each subject participated in the transfer and control conditions (Fig. 1). These two conditions used different keyboard layouts, and different letters in subsets A, B and C. Results for the two conditions were obtained a week apart, with a fully counter-balanced design. The skew of reaction-time (RT) distributions was reduced by taking logarithms (designated as  $RT_{log}$ ). We binned each subject's  $RT_{log}$  values (16 trials per bin in session 1; 8 trials per bin in sessions 2 and 3) and analysed bin means using repeated-measures two-factor MANOVAs (condition and bin number); we then used linear contrasts to test specific hypotheses, denoted  $F_{LC}$ . Response accuracy was not significantly less than 100% in any session. After learning subsets A and B,  $RT_{log}$  for subset B (session 3) was significantly smaller after relearning subset A than after learning subset C (mean reaction times, 0.990 s and 1.123 s, respectively; Fig. 1c).

It is possible that this difference was caused by increased reaction time in the control condition (for example, through 'interference' from learning new items in subset C), rather than by reduced reaction time in the transfer condition. However, two findings are inconsistent with this interpretation. First, comparison of performance with and without learning of new items (that is, testing subset B in the control condition and relearning subset A in the transfer condition) shows no difference ( $F_{LC}(1)=0.821$ ,  $P=0.352$ ). Second, in the



**Figure 1** Reaction time ( $RT_{log}$ ) plotted against trial number for three different experimental sessions (solid lines, transfer condition; dashed lines, control condition; error bars, standard errors of bin means; see text). **a**, Session 1: learning intermixed letter subsets A and B.  $RT_{log}$  decreases as letter positions are learned ( $F(29,290)=17.44$ ,  $P<0.001$ ), with no effect of condition ( $P=0.560$ ) and no condition  $\times$  trial interaction ( $P=0.697$ ). **b**, Session 2: relearning subset A (transfer condition), and learning new subset C (control condition).  $RT_{log}$  during relearning of subset A was less than  $RT_{log}$  during learning of subset C ( $F_{LC}(1)=34.298$ ,  $P<0.001$ ). **c**, Session 3: testing subset B.  $RT_{log}$  for subset B was significantly smaller after relearning subset A (transfer condition) than after learning subset C (control condition) ( $F_{LC}(1)=7.930$ ,  $P=0.006$ ).

transfer condition,  $RT_{log}$  during session 3 was significantly smaller than during session 2 ( $F_{LC}(1)=3.647$ ,  $P=0.036$ ). This suggests that relearning of subset A in session 2 involved implicit relearning of subset B.

This investigation was stimulated by mathematical analyses of neural-network models. Our findings are consistent with a form of spontaneous recovery (the transfer effect) and support a computational account of learning and relearning in human memory. Moreover, the transfer effect may contribute to the savings that are observed when forgotten associations come to be relearned<sup>4</sup>.

**James V. Stone\*, Nicola M. Hunkin†, Angela Hornby\***

\*Department of Psychology and †Academic Neurology Unit, Sheffield University,

## brief communications

Sheffield S10 2TP, UK  
e-mail: j.v.stone@sheffield.ac.uk

1. Hinton, G. E. & Sejnowski, T. J. in *Parallel Distributed Processing* (eds Rumelhart, D. & McClelland, J.) 282–317, (MIT Press, Cambridge, Massachusetts, 1986).
2. Hinton, G. E. & Plaut, D. *Proc. Ninth Annu. Cogn. Sci. Soc.* 177–186 (1987).
3. Harvey, I. & Stone, J. V. *Neural Comp.* 8, 697–704 (1996).
4. Ebbinghaus, H. *Über das Gedächtnis* (Dunker, Leipzig, 1885).

### Parasitic infection

## Hunger tolerance and *Leishmania* in sandflies

The sandfly *Phlebotomus papatasi* transmits *Leishmania major*, the agent of cutaneous leishmaniasis, in desert and savannah regions of the Old World<sup>1</sup>, where seasonal stress of dehydration and heat reduces the quantity of sugar in plant leaves<sup>2</sup>. Without essential sugar, only a few flies that feed on leaves can survive for long enough to deposit eggs and transmit *Leishmania*<sup>3,4</sup>. Accordingly, selection for hunger tolerance may also select for pathogen susceptibility in flies. Here we provide evidence of a link between these advantageous and costly<sup>5,6</sup> properties by testing the susceptibility of flies selected by sugar deprivation and of flies from irrigated and arid habitats.

The experimental flies were obtained from a colony maintained on sucrose solution that was established in 1983 using flies from a desert habitat in the Jordan Valley. Susceptibility to infection with *L. major* promastigotes, the flagellated, parasitic form that inhabits sandflies, was evaluated sporadically using a standard procedure (Table 1)<sup>7</sup>. The previously high infection rate (338 of 394 flies, 85.7%) has declined in recent years to only 27.0% (48 of 178).

Experimental native plants *Atriplex halimus* (Chenopodiaceae) and *Malva nicaeensis* (Malvaceae), and the domestic *Bougainvillea glabra* (Nyctaginaceae), were chosen in a dry area. Sugar levels<sup>4,8</sup> in the leaves of these plants were 3.6–13.8 times lower than in those of irrigated conspecifics. Groups of female *P. papatasi* were each fed exclusively on branches of one of the

plants<sup>4</sup> and other groups were starved until about half of the flies died (4–6 days). The survivors and control groups, one previously maintained for 6 days on 30% sucrose solution and another of day-old flies, were then infected<sup>7</sup>. Significantly more infections, including greater numbers of parasites, were found in flies selected for hunger tolerance compared with control flies (Table 1).

As the increased susceptibility to parasitic infection could have resulted from pre-infection hunger rather than from being selected, we tested day-old progeny of hunger-selected females. The frequency of infection in the progeny series was 64.3% (9 of 14), 79.7% (59 of 74) and 65.6% (21 of 32) (parents fed on *B. glabra* or *M. nicaeensis* or starved, respectively), whereas only 25.4% (45 of 177) of control flies were infected. This difference is statistically as significant as that of the parent generation (Table 1). Similar differences in susceptibility were also seen after progeny of hunger-selected and unselected flies were infected with amastigotes, the mammalian parasite form of *L. major* (results not shown). Assuming that hunger resistance and host competence for *L. major* improve with size, we measured and found no difference between the experimental and control fly groups (results not shown).

We also tested the correlation between the abundance of plant sugar in an oasis (Neot-Hakikar, Israel) and its scarcity at an arid site (in the Jordan Valley), and the susceptibility of the local *P. papatasi* populations to *L. major*. We first tested trapped flies for the presence of sugar in the gut<sup>3,8</sup> — most of the oasis flies (69 of 72, 95.8%), but only one-third of the arid-site flies (60 of 186, 32.3%), were positive. We then infected first-generation progeny of these populations with *L. major* promastigotes<sup>7</sup>. Only 24.8% (25 of 101) of the oasis progeny retained parasites (mean 5 s.e. per fly, 243571 promastigotes) compared with 82.5% (156 of 189) of the progeny of arid-site flies (mean 5 s.e. per fly, 3,412,568 promastigotes). We conclude that the oasis-adapted population is less susceptible to *L. major* than are similar flies from an

arid area.

The cost of parasitic infection seems to be sufficient to elicit natural selection against competent hosts<sup>9</sup>. As far as we know, our findings represent the first evidence that this may be compensated for by linkage to endurance against environmental stress. Seasonal and incidental stress are prevalent in the life of insects<sup>10</sup>, including those that are vectors of disease, and selection may often favour lines of parasites that are adapted to stress-resistant vectors.

Yosef Schleim, Raymond L. Jacobson

Department of Parasitology, Hebrew University  
Hadassah Medical School, PO Box 12272,  
Jerusalem 91120, Israel  
e-mail: schleim@md.huji.ac.il

1. Ashford, R. W. & Bettini, S. in *The Leishmaniasis in Biology and Medicine* Vol. 1 (eds Peters, W. & Killick-Kendrick, R.) 365–424 (Academic, London, 1987).
2. Hopkins, W. G. *Introduction to Plant Physiology* (Wiley, New York, 1995).
3. Schleim, Y. & Jacobson, R. L. *Med. Vet. Entomol.* 13, 65–71 (1999).
4. Schleim, Y. & Jacobson, R. L. *J. Med. Entomol.* 37, 319–324 (2000).
5. Schleim, Y., Jacobson, R. L. & Messer, G. *Proc. Natl Acad. Sci. USA* 89, 9944–9948 (1992).
6. Schaub, G. A. *Parasitol. Today* 10, 463–468 (1994).
7. Schleim, Y., Warburg, A., Schnur, L. F. & Shlomai, J. *Acta Tropica* 40, 65–70 (1983).
8. Van Handel, E. *Anal. Biochem.* 19, 193–194 (1967).
9. Yan, G., Severson, D. W. & Christensen, B. M. *Evolution* 51, 441–450 (1997).
10. Harshman, L. G., Hoffmann, A. A. & Clark, A. G. *J. Evol. Biol.* 12, 370–379 (1999).
11. Zar, J. H. *Biostatistical Analysis* 4th edn (Prentice-Hall, New Jersey, 1999).

### correction

#### Ancient homes for hard-up hermit crabs

David K. A. Barnes

*Nature* 412, 785–786 (2001)

It has been drawn to my attention that I inadvertently omitted to cite previous work relevant to my results — hermit crabs have also been found to use fossils in Bermuda<sup>1,2</sup>, although in a different context. For some *Coenobita clypeatus*, namely those in Bermuda, the only shells available of appropriate size are fossils, and therefore little choice or active selection is being made by the resource users. The hermit crab *Coenobita rugosus* (in Madagascar), like most others elsewhere, is not able to excavate buried shells. Haas<sup>1</sup>, however, has observed *C. clypeatus* digging up and clearing debris from partially buried shells in Bermuda, which was also seen by Kellogg<sup>3</sup> in laboratory conditions.

1. Haas, F. *Ecology* 31, 152 (1950).
2. Walker, S. E. *Palaos* 9, 403–412 (1994).

brief communications is intended to provide a forum for both brief, topical reports of general scientific interest and technical discussion of recently published material of particular interest to non-specialist readers. Priority will be given to contributions that have fewer than 500 words, 10 references and only one figure. Detailed guidelines are available on *Nature's* website ([www.nature.com](http://www.nature.com)) or on request from [nature@nature.com](mailto:nature@nature.com)

**Table 1 Susceptibility of starvation-tolerant sandflies to *Leishmania major***

| Pre-infection diet      | Flies dead | No. of flies | Flies infected | Parasites per fly |      |
|-------------------------|------------|--------------|----------------|-------------------|------|
|                         | (%)        |              | (%)            | Mean              | s.e. |
| <i>Atriplex halimus</i> | 56.7       | 18           | 61.1*          | 4,075             | 434  |
| <i>Malva nicaeensis</i> | 57.6       | 16           | 75.0*          | 588               | 113  |
| Water only              | 53.6       | 33           | 78.8*          | 920               | 243  |
| 30% sucrose             | 0          | 78           | 12.8           | 211               | 136  |
| None                    | 0          | 177          | 25.4†          | 531               | 248  |

Experimental flies survived for 4–6 days when fed on low-sugar plant branches or starved (water alone). One control group received 30% sucrose for 6 days before artificial infection<sup>7</sup>; another was 1 day old when infected. Infective meals consisted of 12.10<sup>6</sup> mi<sup>-1</sup> promastigotes (MHOM/IL/90/LRC-L585 and IPAP/IL/84/LRC-L465) in rabbit blood. Blood-fed flies were dissected and their parasites counted after 6 days of maintenance on 30% sucrose solution and water.

\*Proportional data arcsine-transformed and significantly different from a standard control ( $P < 0.001$ ; Fisher's exact test)<sup>11</sup>.

†Combined results of nine series.



# Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems

D. S. Schimel<sup>1,2†</sup>, J. I. House<sup>1</sup>, K. A. Hibbard<sup>2</sup>, P. Bousquet<sup>3</sup>, P. Ciais<sup>3</sup>, P. Peylin<sup>4</sup>, B. H. Braswell<sup>13</sup>, M. J. Apps<sup>5</sup>, D. Baker<sup>6</sup>, A. Bondeau<sup>7</sup>, J. Canadell<sup>8</sup>, G. Churkina<sup>1</sup>, W. Cramer<sup>7</sup>, A. S. Denning<sup>9</sup>, C. B. Field<sup>10</sup>, P. Friedlingstein<sup>3</sup>, C. Goodale<sup>10</sup>, M. Heimann<sup>1</sup>, R. A. Houghton<sup>11</sup>, J. M. Melillo<sup>12</sup>, B. Moore III<sup>13</sup>, D. Murdiyarso<sup>14</sup>, I. Noble<sup>15</sup>, S. W. Pacala<sup>16</sup>, I. C. Prentice<sup>1</sup>, M. R. Raupach<sup>17</sup>, P. J. Rayner<sup>18</sup>, R. J. Scholes<sup>19</sup>, W. L. Steffen<sup>20</sup> & C. Wirth<sup>1</sup>

**Knowledge of carbon exchange between the atmosphere, land and the oceans is important, given that the terrestrial and marine environments are currently absorbing about half of the carbon dioxide that is emitted by fossil-fuel combustion. This carbon uptake is therefore limiting the extent of atmospheric and climatic change, but its long-term nature remains uncertain. Here we provide an overview of the current state of knowledge of global and regional patterns of carbon exchange by terrestrial ecosystems. Atmospheric carbon dioxide and oxygen data confirm that the terrestrial biosphere was largely neutral with respect to net carbon exchange during the 1980s, but became a net carbon sink in the 1990s. This recent sink can be largely attributed to northern extratropical areas, and is roughly split between North America and Eurasia. Tropical land areas, however, were approximately in balance with respect to carbon exchange, implying a carbon sink that offset emissions due to tropical deforestation. The evolution of the terrestrial carbon sink is largely the result of changes in land use over time, such as regrowth on abandoned agricultural land and fire prevention, in addition to responses to environmental changes, such as longer growing seasons, and fertilization by carbon dioxide and nitrogen. Nevertheless, there remain considerable uncertainties as to the magnitude of the sink in different regions and the contribution of different processes.**

A proportion of the carbon dioxide emitted to the atmosphere by fossil-fuel burning and terrestrial processes (mainly deforestation) is taken up by the oceans and the terrestrial biosphere. High-precision atmospheric observations of concentrations of CO<sub>2</sub> and O<sub>2</sub> (as O<sub>2</sub>:N<sub>2</sub> ratios) make it possible to partition the uptake of atmospheric CO<sub>2</sub> between the land and ocean with increased confidence<sup>1–3</sup>. Global carbon budgets have been updated in the most recent IPCC assessment<sup>4</sup> using this approach (Table 1). The net terrestrial biospheric flux between the land and atmosphere was about –0.2 gigatonnes of carbon per year (Gt C yr<sup>–1</sup>) in the 1980s and –1.4 Gt C yr<sup>–1</sup> in the 1990s (the negative sign denotes a flux from the atmosphere). The net terrestrial biospheric flux is the difference between terrestrial uptake (sinks) and sources. Estimates of land-use change suggest emissions in the range of +0.6 to +2.5 Gt C yr<sup>–1</sup> for the 1980s, largely from deforestation in the tropics<sup>4</sup>. If emissions due to land-use change were of a similar magnitude in the 1990s, this would imply a residual terrestrial sink of between about –2 and –4 Gt C yr<sup>–1</sup>.

## Spatial patterns of carbon uptake

Over the past two decades, evidence has accumulated of significant contributions of extratropical Northern Hemisphere land areas to the global uptake of anthropogenic CO<sub>2</sub>: this evidence has been obtained from analysis of land inventory data<sup>5–8</sup>, atmospheric CO<sub>2</sub> data<sup>9–13</sup>, atmospheric O<sub>2</sub> data<sup>1–3</sup>, isotopic analyses<sup>3,14</sup>, studies of land-use change<sup>15</sup> and ecosystem process models<sup>16,17</sup>. Figure 1 shows zonal flux estimates from an ensemble of global atmospheric inverse model calculations: this is a method that back-calculates

sources and sinks of CO<sub>2</sub> from the distribution of atmospheric concentrations, resulting in a range of estimates of the northern extratropical net land sink from –0.6 to –2.3 Gt C yr<sup>–1</sup> in the 1980s<sup>13</sup>.

Broad longitudinal partitioning of fluxes remains less certain than broad latitudinal partitioning, mainly because the former give rise to smaller differences in concentration<sup>14,18</sup>. Inverse model results are highly sensitive to the subset of atmospheric data used, to the spatial and temporal resolution of the calculation method, and to the atmospheric model used<sup>12</sup>, as indicated in Table 2. For example, the estimates of the North American sink range from 0 to 88% of the total northern land sink, depending on the approach.

Recent inverse modelling studies of the northern extratropical sink do not indicate a large imbalance of fluxes between North

**Table 1 Contemporary carbon budgets for the 1980s and 1990s**

|                                                     | 1980s* (Gt C yr <sup>–1</sup> ) | 1990s* (Gt C yr <sup>–1</sup> ) |
|-----------------------------------------------------|---------------------------------|---------------------------------|
| Emissions (fossil-fuel burning, cement manufacture) | 5.4 ± 0.3                       | 6.3 ± 0.4                       |
| Atmospheric increase                                | 3.3 ± 0.1                       | 3.2 ± 0.1                       |
| Ocean–atmosphere flux                               | –1.9 ± 0.5                      | –1.7 ± 0.5                      |
| Land–atmosphere flux                                | –0.2 ± 0.7                      | –1.4 ± 0.7                      |
| Emissions due to land-use change                    | 1.7 (0.6 to 2.5)†               | Assume 1.6 ± 0.8‡               |
| Residual terrestrial sink                           | –1.9 (–3.8 to 0.3)              | –2 to –4 (Highly uncertain)     |

Negative values denote flux from the atmosphere, that is, ocean or land uptake.

\* From ref. 4.

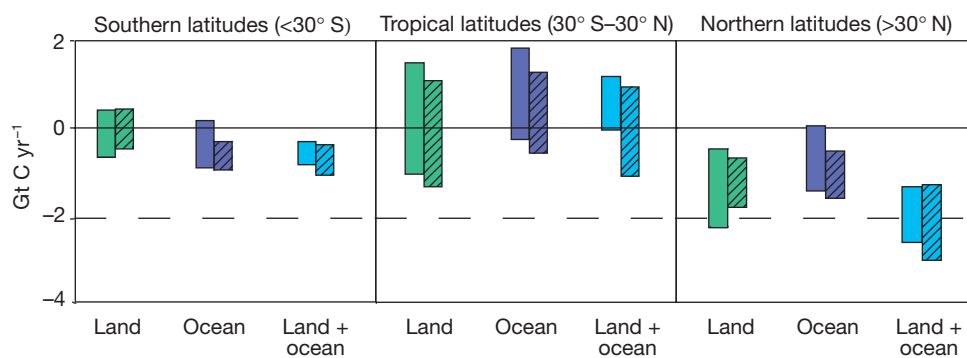
† The range of estimates available to IPCC 2001 (ref. 4).

‡ Based on the early 1990s only and not the full decade (ref. 24).

<sup>1</sup> Max Planck Institute für Biogeochemie, Jena, Germany; <sup>2</sup> IGBP/GAIM, University of New Hampshire, Morse Hall, Durham, New Hampshire 03824, USA; <sup>3</sup> LSCE Unité mixte CEA-CNRS, Bat. 709, CE L'Orme des Merisiers, 91191, Gif sur Yvette France; <sup>4</sup> Laboratoire de Biogéochimie Isotopique, Unité mixte CNRS-UPMC-INRA, 4 Place Jussieu, 75252 Paris, France; <sup>5</sup> Natural Resources Canada, Canadian Forest Service Northern Forestry Center, 5320 122 Street, Edmonton, Alberta, Canada; <sup>6</sup> NCAR, 1850 Table Mesa Drive, Boulder 80303, USA; <sup>7</sup> Potsdam Institute for Climate Impact Research, Telegrafenberg C4, 14473 Potsdam; <sup>8</sup> GCTE International Project Office, CSIRO Sustainable Ecosystems, PO Box 284, Canberra, ACT 2601, Australia; <sup>9</sup> Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado 80523-1371, USA; <sup>10</sup> Carnegie Institution of Washington, Department of Plant Biology, 260 Panama Street, Stanford, California 94305, USA; <sup>11</sup> Woods Hole Research Center, PO Box 296, Woods Hole, Massachusetts 02543, USA; <sup>12</sup> Ecosystems Center, Marine Biological Laboratory, Woods Hole,

Massachusetts 02543, USA; <sup>13</sup> University of New Hampshire, Institute for the Study of Earth, Oceans and Space, Durham, New Hampshire 03824, USA; <sup>14</sup> GCTE Impacts Center for Southeast Asia, Jalan Raya Tajur Km 6, POB 116, Bogor, Indonesia; <sup>15</sup> Ecosystem Dynamics, RSBS, Australian National University, POB4 Acton, Canberra, ACT 0200, Australia; <sup>16</sup> Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Hampshire 08544-1003, USA; <sup>17</sup> CSIRO Division of Land and Water, GPO Box 1666, Canberra ACT 2601, Australia; <sup>18</sup> CSIRO-DAR, PMB #1, 3195 Aspendale, Australia; <sup>19</sup> Environmentek, CSIR, PO Box 395, Pretoria 0001, South Africa; <sup>20</sup> IGBP Secretariat, Box 50005, S104-05, Stockholm, Sweden.

<sup>21</sup> Present address: National Center for Atmospheric Research, 1850 Table Mesa Drive, Boulder, Colorado 80305, USA.



**Figure 1** Zonal distribution of terrestrial and oceanic carbon fluxes. These data were deduced from eight inverse models using different techniques and sets of atmospheric observations after accounting for fossil-fuel emissions (not shown)<sup>13</sup>. Results are shown for the 1980s (plain bars) and for 1990–96 (hatched bars). The bars indicate

the full range of results from the models. Positive numbers are fluxes to the atmosphere. Note that data and time-intervals are not precisely consistent with those used in other analyses presented in this Progress Article, and so the global totals may differ from those mentioned elsewhere.

America and Eurasia<sup>3,11,12,19–21</sup>, in contrast to earlier reports of a dominant sink in North America<sup>22</sup>. Given the controversy over the east–west partitioning of sinks, we need to consider if there is a biophysical or climatic reason to expect North America and Eurasia to differ greatly in ecosystem carbon uptake. Although the mean of the results in Table 2 suggests that Eurasia is a sink about twice the size of North America, on a per unit land area basis, the mean estimated uptake rates on North America and Eurasia are broadly similar at  $-32$  and  $-39 \text{ g C m}^{-2} \text{ yr}^{-1}$  respectively in the early 1990s (note that global uptake was high in this period relative to the rest of the decade, Fig. 2). On a vegetated area basis<sup>23</sup>—that is, after excluding ‘bare’ areas such as deserts and ice—estimated uptake is still similar for both continents. Major climate differences between the two regions can be normalized by weighting the areas by growing season length, thereby providing an integrative index of ecosystem activity (Table 2): this again gives similar rates of uptake (per growing season day times area) on each continent. Therefore uptake rates based on mean estimates from the inversions in Table 2

are not inconsistent with expectations based on total area or bioclimatically weighted area on the two continents. New inverse modelling intercomparisons<sup>21</sup> and ground-based studies<sup>8</sup> are consistent with the breakdown presented here. Given the extreme uncertainty in partitioning of fluxes (Table 2), the partitioning cannot yet be regarded as definitive.

In the tropics, atmospheric inverse model calculations do not detect the large  $\text{CO}_2$  source that would be expected from deforestation alone, but show variable results clustering around zero. This seems to imply the existence of a sink that balances the deforestation source<sup>12,13</sup>, although results are poorly constrained by the sparse measurements in the tropics. If the deforestation source is  $+1.6 \text{ Gt C yr}^{-1}$  for the first half of the 1990s<sup>24</sup>, then a net sink of  $-0.4 \text{ Gt C yr}^{-1}$  (Table 2) implies an uptake of  $-2.0 \text{ Gt C yr}^{-1}$  by tropical ecosystems. Local studies show carbon uptake in a range of mature tropical forest types<sup>25,26</sup>, but it is not possible to extrapolate these to the entire tropical region due to high heterogeneity of tropical ecosystems. Because of sparse atmospheric and ecological

**Table 2** Estimated distribution of net land–atmosphere carbon fluxes between North America and Eurasia

|                                       | Net land–atmosphere flux, NBP† (Gt C yr <sup>-1</sup> ) | Percentage of Northern Hemisphere uptake | Total land area* (10 <sup>12</sup> m <sup>2</sup> ) | Net flux per unit area (g C m <sup>-2</sup> yr <sup>-1</sup> ) | Vegetated area† (10 <sup>12</sup> m <sup>2</sup> ) | Net flux per unit vegetated area (g C m <sup>-2</sup> yr <sup>-1</sup> ) | Growing-season-weighted (GSW) area‡ (10 <sup>15</sup> m <sup>2</sup> d) | Net flux per unit GSW area (g C m <sup>-2</sup> d <sup>-1</sup> ) | Modelled NPP§ (Gt C yr <sup>-1</sup> ) | Net flux (~NBP) per unit NPP |
|---------------------------------------|---------------------------------------------------------|------------------------------------------|-----------------------------------------------------|----------------------------------------------------------------|----------------------------------------------------|--------------------------------------------------------------------------|-------------------------------------------------------------------------|-------------------------------------------------------------------|----------------------------------------|------------------------------|
| North America                         | -0.8<br>(-2.1 to +0.1)                                  | 0 to 88%                                 | 25                                                  | -32                                                            | 20                                                 | -40                                                                      | 3.1                                                                     | -0.26                                                             | 5 to 9                                 | 9 to 16%                     |
| Eurasia                               | -1.7<br>(-2.5 to -0.2)                                  | 13 to 100%                               | 42                                                  | -39                                                            | 36                                                 | -46                                                                      | 5.2                                                                     | -0.33                                                             | 8 to 15                                | 11 to 21%                    |
| Northern extratropical total          | -2.4<br>(-4.3 to -1.5)                                  | NA                                       | 67                                                  | -36                                                            | 56                                                 | -44                                                                      | 8.2                                                                     | -0.30                                                             | 13 to 23                               | 11 to 18%                    |
| Tropical and southern temperate total | -0.4<br>(-1.2 to +0.8)                                  | NA                                       | 70                                                  | -5¶                                                            | 47                                                 | -8¶                                                                      | 15.6                                                                    | -0.03¶                                                            | 28 to 45                               | ~1%¶                         |
| Global total                          | -2.8<br>(-4.3 to -1.7)                                  | NA                                       | 137                                                 | -20                                                            | 104                                                | -27                                                                      | 23.8                                                                    | -0.12                                                             | 42 to 68                               | 4 to 6%                      |

A negative sign denotes a flux from the atmosphere into ecosystems. Data here are for 1990–94 from an ensemble of inverse experiments<sup>12</sup>. The net land–atmosphere flux is the balance of terrestrial biospheric sources and sinks. Inverse modelling techniques attempt to resolve sources and sinks of carbon by back-calculating from measurements of atmospheric  $\text{CO}_2$ . In this experiment<sup>12</sup>, three atmospheric transport models were used, with three levels of spatial resolution (7, 12 and 17 land and ocean regions) and three temporal averaging approaches (annual average data with annual fluxes, monthly data with annual fluxes, and monthly data with monthly fluxes). Values shown are mean annual fluxes and the full range of model results in parentheses. The ranges indicate the high sensitivity of the analyses to the techniques used. The estimates are during a time period 1990–94 when the global mean land uptake was roughly 30% higher than the long-term mean (Fig. 2), thus the sinks estimated are high compared with the decadal average in Table 1, but the relative distribution between the northern continents should be roughly independent of time period. Additionally, fluxes calculated in this analysis have not been adjusted to account for river transport which will show up as a larger land sink. The tropics are shown here for comparison, but results are more uncertain due to the sparsity of the atmospheric measurement network. NA, not applicable.

\* The land area relates to the regions covered in the inverse modelling analyses<sup>12</sup> and excludes Greenland.

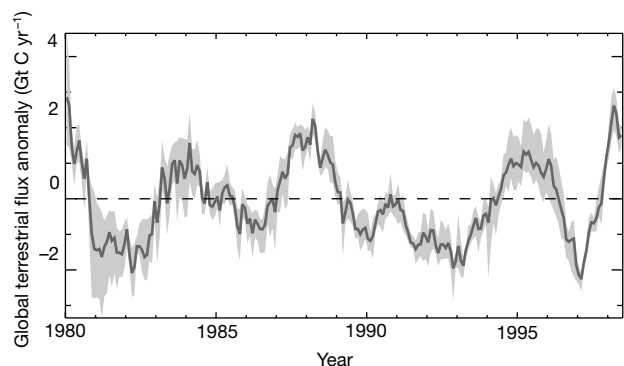
† The vegetated area is estimated from a satellite-based land cover classification<sup>23</sup> in coincidence with the inverse model regions, and excluding land classified as ‘bare’ (that is, deserts and ice).

‡ Long-term average growing season length — calculated from the satellite normalized difference vegetation index (NDVI) from the NOAA AVHRR instrument — is the period (days) when the NDVI is above a threshold value (0.3). While the NDVI signal comes from the plant canopy, the factors that control plant growing season length also affect soil processes. The growing season length was computed for each cell in the NDVI data set, multiplied by cell area and the products summed to produce a growing season weighted area.

§ The modelled NPP values are estimates of long-term average NPP and are the highest and lowest estimates from five ecosystem models SLAVE<sup>42</sup>, CASA, GLO-PEM, Biome-BGC, and TEM<sup>47</sup>.

¶ As inverse atmospheric methods detect the net effect on the atmosphere of all sources, sinks and removals they can be considered to approximate net biome productivity (NBP) over land (NBP is NPP minus losses due to death and decomposition, fire, and removals/lateral transfers).

¶¶ The tropical flux is influenced by the large deforestation emissions, leading to a lower net flux per unit land area and NBP to NPP ratio than in northern extratropical regions.



**Figure 2** An illustrative plot of the interannual variability of global terrestrial carbon exchange, as deduced using inverse modelling<sup>10</sup>. The anomalies plotted on the y-axis are deviations from the long-term mean flux. The black line is the average of eight sensitivity inversions using the same model, and shaded areas represent the range of values from the inversions. Positive numbers are fluxes to the atmosphere. Although this analysis is from a single model, other analyses produce a similar pattern, although with variation in the details of the magnitude and phasing of the fluxes<sup>29</sup>. Most current analyses suggest that the land biosphere plays a larger role than the oceans in the global interannual variability of atmospheric CO<sub>2</sub> (refs 3, 10, 29).

sampling, and complex meteorology, estimates of tropical fluxes have high uncertainty.

### Interannual variability

The year-to-year variability of the average annual growth in atmospheric CO<sub>2</sub> concentrations is high<sup>27,28</sup>. Top-down atmospheric calculations, eddy covariance flux observations, long-term ecosystem carbon budgeting studies and modelling all indicate large year-to-year variability in terrestrial metabolism<sup>10,17,29–36</sup> (Fig. 2). Long-term processes, such as increasing CO<sub>2</sub> concentration and land-use changes, are the main drivers of the mean fluxes, but probably have a small effect on year-to-year variations<sup>37</sup>. These terrestrial flux variations are probably caused by the effect of climate on carbon pools with short lifetimes (foliage, plant litter, soil microbes) through variations in photosynthesis, respiration, nutrient cycling and fire.

The net terrestrial sink appears to have increased on average from the 1980s to the 1990s (Table 1). The contribution of different processes remains difficult to quantify, but the unusually large sink in the early 1990s is thought to be largely a consequence of climate variability rather than a response to a systematic trend. Globally, there appears to be a net release of carbon to the atmosphere during warm and dry years, and a net uptake during cooler years<sup>30,38</sup>. Recently, evidence for links between the El Niño/Southern Oscillation cycle and atmospheric CO<sub>2</sub> have become stronger<sup>29,30,39</sup>.

### Controls over terrestrial carbon exchange

The similarity of fluxes of carbon in North America and Eurasia per unit land area or bioclimatic index does not suggest an asymmetry between uptake processes on the two continents, in contrast to factors that may indicate spatial variability of processes. For example, nitrogen deposition is 2 to 4 times higher in Europe compared to the United States<sup>40</sup>, and land management practices and history differ. In the United States, studies indicate that much (possibly most) of the sink is due to changing land use (including abandonment of surplus agricultural land) and more subtle management effects, such as reduced fire frequency leading to woody encroachment<sup>15–17,41,42</sup>. European studies show large effects of both land use/management changes and increased tree growth

possibly due to CO<sub>2</sub> fertilization and N deposition<sup>6,17,43–45</sup>. Chinese inventory studies also show a significant carbon sink resulting from afforestation and reforestation programs<sup>46</sup>.

Combining the atmospherically derived fluxes presented in Table 2 (approximately equivalent to net biome productivity, NBP) with model-derived long-term averages of net primary productivity (NPP)<sup>43,47</sup> suggests that 10–20% of carbon annually fixed by plants in the northern extratropics in the early 1990s remained in the biosphere (Table 2). This can be compared to inventory-based estimates of NPP to NBP ratios over large regions: for example, Nilsson *et al.*<sup>48</sup> estimated this ratio in Russia to be 5% in forests, 10% in wetlands and 16% in grass/shrublands, while Schulze *et al.*<sup>49</sup> estimate it as nearly 25% in managed European forests. All of these estimates suggest carbon accumulation in ecosystems recovering from disturbance or undergoing intensive management and, as such, are high compared with chronosequence studies that suggest much lower long-term values in natural landscapes<sup>50</sup>.

In some regions, long-term climate changes—such as enhanced growing season in high northern latitudes and increased drought in the tropics—may be causing trends in terrestrial ecosystem processes<sup>17,51,52</sup>. Much of Siberia has been warming at ~0.5 °C per decade during 1960–2000, and increased water stress has been documented in Alaska<sup>53,54</sup>. Recent increases in wildfire and insects appear to have converted Siberian and North American boreal forests into carbon sources<sup>48,54,55</sup>, although the northern extratropical regions as a whole are a sink. The magnitude and contribution of different processes to tropical sinks remain largely unknown, although modelling studies suggest that there may be a significant component due to increasing atmospheric CO<sub>2</sub> concentrations<sup>17</sup>.

Sinks of today's magnitude cannot be counted on to operate steadily into the future, as all of the key processes are likely to diminish. Uptake due to land-use change such as forest growth on agricultural land will eventually decline as the forests reach maturity. Fertilization by both CO<sub>2</sub> and N is expected to saturate at high levels, and as other resources become limiting<sup>56</sup>. The effects of climate change on ecosystems is expected to reduce sinks at a global scale<sup>17,52</sup>. The net terrestrial sink may thus disappear altogether in the future, although model projections of these dynamics differ greatly<sup>4,47,52</sup>.

### Central issues for future research

The response of ecosystems to regionally heterogeneous stimuli such as historical land use, nitrogen deposition rate, or rainfall and temperature anomalies and their effects on carbon fluxes are not predictable from global averages. Documenting these fluxes and elucidating the processes controlling them will require both regional observing systems and improved historical data on management and natural disturbances (fire, insects, and so on). Spatially extensive observations that are temporally repeatable—either inventories or a combination of inventory and eddy covariance data—are required. Uncertainty in regional and global fluxes could be reduced by improving the distribution of atmospheric CO<sub>2</sub> measurements over the continents and tropics: this could be done by enhancing the existing atmospheric flask sampling networks<sup>28</sup> and by using aircraft or tall towers for sampling. In addition, increased precision in measuring CO<sub>2</sub> concentrations from satellites, when it becomes feasible, could provide globally comprehensive observations.

Emissions from land-use changes account for a large (approximately 20%) but uncertain fraction of anthropogenic CO<sub>2</sub> emissions<sup>4</sup>. The lack of a clear atmospheric signal of tropical deforestation, and the implied large tropical sink, causes a major uncertainty in our ability to balance the terrestrial carbon cycle. The current lack of adequate data on land-use change and ecosystem processes for the tropics makes it impossible to evaluate ground-based versus atmospheric estimates as is done for the northern extratropics.



Atmospheric inverse calculations simultaneously derive ocean and land fluxes, and so uncertainty in terrestrial and marine fluxes are not independent. Understanding and stabilization of the Earth system will require these uncertainties to be addressed by efforts linking terrestrial ecology with the other Earth sciences. □

1. Bender, M., Ellis, T., Tans, P., Francey, R. & Lowe, D. Variability in the O<sub>2</sub>/N<sub>2</sub> ratio of southern hemisphere air, 1991–1994: Implications for the carbon cycle. *Glob. Biogeochem. Cycles* **10**, 9–21 (1996).
2. Keeling, R. F., Piper, S. C. & Heimann, M. Global and hemispheric CO<sub>2</sub> sinks deduced from changes in atmospheric O<sub>2</sub> concentration. *Nature* **381**, 218–221 (1996).
3. Rayner, P. J., Enting, I. G., Francey, R. J. & Langenfelds, R. Reconstructing the recent carbon cycle from atmospheric CO<sub>2</sub>, δ<sup>13</sup>C and O<sub>2</sub>/N<sub>2</sub> observations. *Tellus B* **51**, 213–232 (1999).
4. Prentice, I. C. *et al.* in *Climate Change 2001: The Scientific Basis* (eds Houghton, J. T. & Yihui, D.) Ch. 3, 183–237 (Cambridge Univ. Press, Cambridge, 2001).
5. Brown, S. in *Climate Change 1995—Impacts, Adaptations and Mitigation of Climate Change: Scientific-Technical Analyses* (eds Watson, R. T., Zinyowera, M. C., Moss, R. H. & Dokken, D. J.) 773–797 (Cambridge Univ. Press, Cambridge, 1996).
6. Spiecker, H., Mielikäinen, K., Köhl, M. & Skovsgaard, J. P. *Growth Trends in European Forests* (Springer, Berlin, 1996).
7. Brown, S. L. & Schroeder, P. E. Spatial patterns of aboveground production and mortality of woody biomass for eastern US forests. *Ecol. Appl.* **9**, 968–980 (1999).
8. Pacala, S. W. *et al.* Consistent land- and atmosphere-based US carbon sink estimates. *Science* **292**, 2316–2320 (2001).
9. Tans, P. P., Fung, I. Y. & Takahashi, T. Observational constraints on the global atmospheric CO<sub>2</sub> budget. *Science* **247**, 1431–1438 (1990).
10. Bousquet, P. *et al.* Regional changes in carbon dioxide fluxes of land and oceans since 1980. *Science* **290**, 1342–1346 (2000).
11. Baker, D. *Sources and Sinks of Atmospheric CO<sub>2</sub> Estimated from Batch Least-Squares Inversions of CO<sub>2</sub> Concentration Measurements*. Thesis, Princeton Univ. (2000).
12. Peylin, P., Baker, D., Sarmiento, J., Ciais, P. & Bousquet, P. Influence of transport uncertainty on annual mean versus seasonal inversion of atmospheric CO<sub>2</sub> data. *J. Geophys. Res. D* (submitted).
13. Heimann, M. *Atmospheric Inversion Calculations Performed for IPCC Third Assessment Report Chapter 3 (The Carbon Cycle and Atmospheric CO<sub>2</sub>)* (Max-Planck-Institute für Biogeochemie, Jena, 2001).
14. Ciais, P., Tans, P. P., Trolier, M., White, J. W. C. & Francey, R. J. A large northern-hemisphere terrestrial CO<sub>2</sub> sink indicated by the <sup>13</sup>C/<sup>12</sup>C ratio of atmospheric CO<sub>2</sub>. *Science* **269**, 1098–1102 (1995).
15. Houghton, R. A., Hackler, J. L. & Lawrence, K. T. The US carbon budget: Contributions from land-use change. *Science* **285**, 574–578 (1999).
16. Schimel, D. *et al.* Contribution of increasing CO<sub>2</sub> and climate to carbon storage by ecosystems in the United States. *Science* **287**, 2004–2006 (2000).
17. McGuire, A. D. *et al.* Carbon balance of the terrestrial biosphere in the twentieth century: Analyses of CO<sub>2</sub> climate and land-use effects with four process-based ecosystem models. *Glob. Biogeochem. Cycles* **15**, 183–206 (2001).
18. Denning, A. S. *et al.* Three-dimensional transport and concentration of SF<sub>6</sub>—A model intercomparison study (TransCom 2). *Tellus B* **51**, 266–297 (1999).
19. Bousquet, P., Ciais, P., Peylin, P., Ramonet, M. & Monfray, P. Inverse modeling of annual atmospheric CO<sub>2</sub> sources and sinks 1. Method and control inversion. *J. Geophys. Res. D* **104**, 26161–26178 (1999).
20. Kaminski, T., Heimann, M. & Giering, R. A coarse grid three-dimensional global inverse model of the atmospheric transport—1. Adjoint model and Jacobian matrix. *J. Geophys. Res. D* **104**, 18535–18553 (1999).
21. Gurney, K. R., Rayner, P., Law, R. & Denning, S. Atmospheric carbon budget inversion intercomparison: preliminary results. *Trans. Am. Geophys. Union* **81**, 276 (2000).
22. Fan, S. *et al.* A large terrestrial carbon sink in North America implied by atmospheric and oceanic carbon dioxide data and models. *Science* **282**, 442–446 (1998).
23. DeFries, R. S. *et al.* Mapping the land surface for global atmosphere-biosphere models: toward continuous distributions of vegetation's functional properties. *J. Geophys. Res. D* **100**, 20867–20882 (1995).
24. Houghton, R. A. A new estimate of global sources and sinks of carbon from land-use change. *Eos* **81**, S281 (2000).
25. Malhi, Y. *et al.* Carbon dioxide transfer over a Central Amazonian rain forest. *J. Geophys. Res. D* **103**, 31593–31612 (1998).
26. Phillips, O. L. *et al.* Changes in the carbon balance of tropical forests: Evidence from long-term plots. *Science* **282**, 439–442 (1998).
27. Keeling, C. D. & Whorf, T. P. in *Trends: A Compendium of Data on Global Change* (Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, 2001).
28. Conway, T. J., Tans, P. P., Waterman, L. S. & Thoning, K. W. Evidence for interannual variability of the carbon-cycle from the National Oceanic and Atmospheric Administration/Climate Monitoring and Diagnostics Laboratory global air-sampling network. *J. Geophys. Res. D* **99**, 22831–22855 (1994).
29. Rayner, P. J., Law, R. M. & Dargaville, R. The relationship between tropical CO<sub>2</sub> fluxes and the El Niño–Southern Oscillation. *Geophys. Res. Lett.* **26**, 493–496 (1999).
30. Vukicevic, T., Braswell, B. H. & Schimel, D. A diagnostic study of temperature controls on global terrestrial carbon exchange. *Tellus B* **53**, 150–170 (2001).
31. Goulden, M. L., Munger, J. W., Fan, S. M., Daube, B. C. & Wofsy, S. C. Exchange of carbon dioxide by a deciduous forest: Response to interannual climate variability. *Science* **271**, 1576–1578 (1996).
32. Kindermann, J., Würth, G., Kohlmaier, G. H. & Badeck, F. W. Interannual variation of carbon exchange fluxes in terrestrial ecosystems. *Glob. Biogeochem. Cycles* **10**, 737–755 (1996).
33. Parton, W. J. *et al.* Observations and modeling of biomass and soil organic-matter dynamics for the grassland biome worldwide. *Glob. Biogeochem. Cycles* **7**, 785–809 (1993).
34. Kelly, R. H. *et al.* Intra-annual and interannual variability of ecosystem processes in shortgrass steppe. *J. Geophys. Res. D* **105**, 20093–20100 (2000).
35. Prentice, I. C., Heimann, M. & Sitch, S. The carbon balance of the terrestrial biosphere: Ecosystem models and atmospheric observations. *Ecol. Appl.* **10**, 1553–1573 (2000).
36. Knapp, A. K. & Smith, M. D. Variation among biomes in temporal dynamics of aboveground primary production. *Science* **291**, 481–484 (2001).
37. Houghton, R. A. Interannual variability in the global carbon cycle. *J. Geophys. Res. D* **105**, 20121–20130 (2000).
38. Braswell, B. H., Schimel, D. S., Linder, E. & Moore, B. The response of global terrestrial ecosystems to interannual temperature variability. *Science* **278**, 870–872 (1997).
39. Yang, X. & Wang, M. X. Monsoon ecosystems control on atmospheric CO<sub>2</sub> interannual variability: inferred from a significant positive correlation between year-to-year changes in land precipitation and atmospheric CO<sub>2</sub> growth rate. *Geophys. Res. Lett.* **27**, 1671–1674 (2000).
40. Holland, E. A., Dentener, F. J., Braswell, B. H. & Sulzman, J. M. Contemporary and pre-industrial global reactive nitrogen budgets. *Biogeochemistry* **46**, 7–43 (1999).
41. Archer, S., Boutton, T. W. & Hibbard, K. A. in *Global Biogeochemical Cycles and their Interrelationship with Climate* (eds Schulze, E. D.) 115–137 (Academic, London, 2001).
42. Casperson, J. *et al.* Carbon accumulation in U.S. forests is caused overwhelmingly by changes in land use rather than CO<sub>2</sub> or N fertilization or climate change. *Science* **290**, 1148–1151 (2001).
43. Friedlingstein, P. *et al.* On the contribution of the biospheric CO<sub>2</sub> fertilization to the missing sink. *Glob. Biogeochem. Cycles* **9**, 541–556 (1995).
44. Kauppi, P. E., Mielikäinen, K. & Kuusela, K. Biomass and carbon budget of European forests, 1971 to 1990. *Science* **256**, 70–74 (1992).
45. Valentini, R. *et al.* *Accounting for Carbon Sinks in the Biosphere, European Perspective* (CARBOEUROPE European Office, Jena, 2000).
46. Fang, J. Y., Chen, A. P., Peng, C. H., Zhao, S. Q. & Ci, L. Changes in forest biomass carbon storage in China between 1949 and 1998. *Science* **292**, 2320–2322 (2001).
47. Cramer, W. *et al.* Comparing global models of terrestrial net primary productivity (NPP): overview and key results. *Glob. Change Biol.* **5**, 1–15 (1999).
48. Nilsson, S. *et al.* *Full Carbon Account for Russia* (International Institute for Applied Systems Analysis, Laxenburg, 2000).
49. Schulze, E. D. *et al.* in *Carbon and Nitrogen Cycling in European Forest Ecosystems* (ed. Schulze, E. D.) 468–491 (Springer, Berlin, 2000).
50. Harden, J. W., Sundquist, E. T., Stallard, R. F. & Mark, R. K. Dynamics of soil carbon during deglaciation of the Laurentide ice-sheet. *Science* **258**, 1921–1924 (1992).
51. Keyser, A. R., Kimball, J. S., Nemani, R. R. & Running, S. W. Simulating the effects of climate change on the carbon balance of North American high-latitude forests. *Glob. Change Biol.* **6**, 185–195 (2000).
52. Cramer, W. *et al.* Global response of terrestrial ecosystem structure and function to CO<sub>2</sub> and climate change: results from six dynamic global vegetation models. *Glob. Change Biol.* **7**, 357–373 (2001).
53. Jones, P. D., New, M., Parker, D. E., Martin, S. & Rigor, I. G. Surface air temperature and its changes over the past 150 years. *Rev. Geophys.* **37**, 173–199 (1999).
54. Barber, V. A., Juday, G. P. & Finney, B. P. Reduced growth of Alaskan white spruce in the twentieth century from temperature-induced drought stress. *Nature* **405**, 668–673 (2000).
55. Kurz, W. A. & Apps, M. J. A 70-year retrospective analysis of carbon fluxes in the Canadian forest sector. *Ecol. Appl.* **9**, 526–547 (1999).
56. Field, C. B., Chapin, F. S., Matson, P. A. & Mooney, H. A. Responses of terrestrial ecosystems to the changing atmosphere—a resource-based approach. *Annu. Rev. Ecol. Syst.* **23**, 201–235 (1992).

## Acknowledgements

This Progress Article arose from the International Workshop on Integrated Carbon Research and Observations, held on 16–20 October 2000. This workshop was organized under the auspices of the International Geosphere–Biosphere Programme (IGBP) and the Global Terrestrial Observing System (GTOS), and was co-funded by NOAA, the NSF, NASA, the US Department of Energy, the US Department of Interior/US Geological Survey, the US Department of Agriculture and the International Geosphere Biosphere Programme. D.S.S. was also supported by the European Commission, NASA, and the Bundesministerium für Bildung und Forschung.

Correspondence and requests for materials should be addressed to D.S.S. (email: schimel@ucar.edu).

# Genetic tracing reveals a stereotyped sensory map in the olfactory cortex

Zhihua Zou<sup>\*†</sup>, Lisa F. Horowitz<sup>\*†</sup>, Jean-Pierre Montmayeur<sup>\*</sup>, Scott Snapper<sup>‡</sup> & Linda B. Buck<sup>\*</sup>

<sup>\*</sup> Howard Hughes Medical Institute, Department of Neurobiology, Harvard Medical School, Boston, Massachusetts 02115, USA

<sup>‡</sup> Gastrointestinal Unit (Medical Services), Massachusetts General Hospital, Department of Medicine, Harvard Medical School, Boston, Massachusetts 02114, USA

<sup>†</sup> These authors contributed equally to this work

**The olfactory system translates myriad chemical structures into diverse odour perceptions. To gain insight into how this is accomplished, we prepared mice that coexpressed a transneuronal tracer with only one of about 1,000 different odorant receptors. The tracer travelled from nasal neurons expressing that receptor to the olfactory bulb and then to the olfactory cortex, allowing visualization of cortical neurons that receive input from a particular odorant receptor. These studies revealed a stereotyped sensory map in the olfactory cortex in which signals from a particular receptor are targeted to specific clusters of neurons. Inputs from different receptors overlap spatially and could be combined in single neurons, potentially allowing for an integration of the components of an odorant's combinatorial receptor code. Signals from the same receptor are targeted to multiple olfactory cortical areas, permitting the parallel, and perhaps differential, processing of inputs from a single receptor before delivery to the neocortex and limbic system.**

Mammals perceive an immense variety of volatile chemicals as having different odours<sup>1–3</sup>. The initial event in olfactory perception is the detection of odorants by odorant receptors (ORs) on olfactory sensory neurons in the nose<sup>4,5</sup>. Signals generated by these neurons in response to odorants are relayed through the olfactory bulb of the brain, and then the olfactory cortex, ultimately reaching higher cortical areas that are thought to be important in odour discrimination, as well as limbic areas believed to mediate the emotional and physiological effects of odours<sup>1,2,6–9</sup>.

In the olfactory epithelium of the mouse nose there are about 5,000,000 olfactory sensory neurons<sup>10</sup>, each of which expresses only one of about 1,000 different OR genes<sup>11</sup>. Each OR recognizes multiple odorants (refs 11, 46–48), but different odorants are recognized, and thereby encoded, by different combinations of ORs<sup>11</sup>. The olfactory epithelium has four spatial zones that express non-overlapping sets of ORs and project axons to four corresponding zones in the olfactory bulb<sup>10,12–15</sup>. In the nose, neurons expressing the same OR are scattered throughout one zone<sup>10,12</sup>; however, in the olfactory bulb their axons converge at two fixed locations, where they form synapses with bulb mitral and tufted relay neurons in only a few of the bulb's 2,000 glomeruli<sup>13–15</sup>. The result is a stereotyped spatial map in which inputs from different ORs are targeted to different glomeruli and bulb neurons. Consistent with patterns of odour-induced activity<sup>16–22</sup>, an odorant's receptor code is represented in the nose by a dispersed ensemble of neurons and in the bulb by a specific combination of glomeruli.

How is olfactory information organized in the olfactory cortex to ultimately yield the perception of different odours? Anatomical studies indicate that the bulb map is not recapitulated in the olfactory cortex<sup>6–9</sup>. Individual bulb relay neurons, whose axons form synapses in the olfactory cortex, are retrogradely labelled by tracers that are deposited in more than one cortical region<sup>23</sup>, and neurons in different parts of the bulb are retrogradely labelled from a single cortical region<sup>24,25</sup>. These findings have prompted speculation that there may be little or no topographical organization of sensory information in the olfactory cortex. However, a few studies suggest that neighbouring bulb neurons may form synapses at the same sites, hinting that cortical inputs may not be organized diffusely, but in some manner not yet discerned<sup>26,27</sup>.

We were interested in three questions. First, is input from a given OR targeted to discrete sites in the olfactory cortex, as in the bulb; scattered, as in the nose; or arranged in yet another way? Second, do

different areas of the olfactory cortex, which may have different functions, receive inputs from different subsets of ORs, or does each area receive input from the entire OR repertoire? And finally, given that each odorant is recognized by multiple ORs, are inputs from different receptors, which are segregated in the nose and bulb, combined in individual cortical neurons?

## Preparation of mice expressing a tracer with one OR gene

To address these three questions, we used a transneuronal tracer—barley lectin (BL)—that we previously found could cross multiple synapses if expressed from a transgene in mice<sup>28</sup>. When expressed in all olfactory sensory neurons, BL (or wheat germ agglutinin<sup>29</sup>) is transneuronally transferred to second-order bulb neurons and then from those neurons to third-order neurons in the olfactory cortex<sup>28,29</sup>.

To determine how inputs from a single OR are organized in the cortex, we prepared mice in which BL was coexpressed with only one OR gene. Using gene targeting in embryonic stem cells, we modified a single OR gene by inserting—just after its coding region—an internal ribosome entry site (IRES) sequence<sup>15,30</sup> followed by a BL complementary DNA (Fig. 1). The altered stem cells were then used to generate mice that produce a bicistronic messenger RNA from which the OR and BL are independently translated. We did this for two OR genes: *M5*, which is expressed in the most dorsal nasal zone (zone 1), and *M50*, which is expressed in the most ventrolateral zone (zone 4)<sup>10,31</sup>. We then used immunostaining of serial nose and brain sections from these 'M5BL' and 'M50BL' knock-in mice to visualize neurons that contained BL.

In transgenic mice expressing BL under the control of the olfactory marker protein promoter (OmpBL mice), BL is expressed in all olfactory sensory neurons<sup>28</sup>. In those mice, BL labelling is seen in neurons throughout the olfactory epithelium and in the axons of those neurons in all of the glomeruli of the olfactory bulb. In contrast, in M5BL and M50BL knock-in mice, BL was detected in only a small fraction of olfactory sensory neurons in the nose and in only a few glomeruli in the olfactory bulb (Fig. 2). No BL labelling was seen in the nose or bulb of wild-type control mice.

In the nose, labelled neurons were located in zone 1 in M5BL mice ( $n = 10$ ) and in zone 4 in M50BL mice ( $n = 10$ ) (Fig. 2a, b). This patterning resembled that seen in wild-type mice by *in situ* hybridization with *M5* and *M50* antisense RNA probes (Fig. 2c, d)<sup>10,28</sup>. The BL<sup>+</sup> glomeruli of the knock-in mice ( $n = 10$  per knock-in strain)

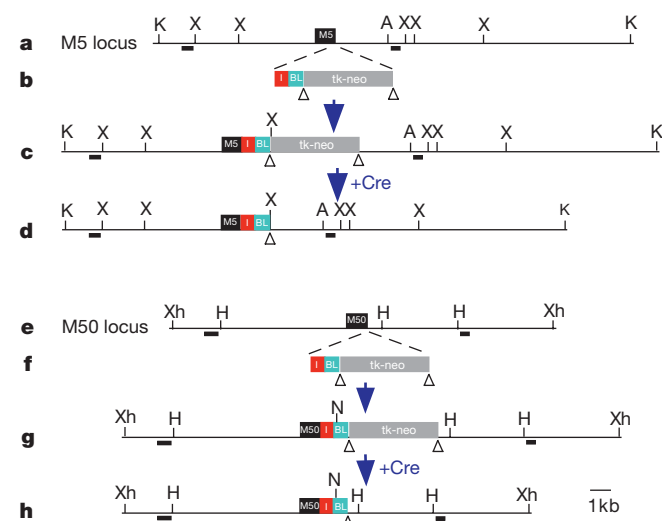
were also similar in both number and position to glomeruli that hybridize to M5 and M50 probes in wild-type animals (Fig. 2e, f, h, i)<sup>13,28</sup>. Moreover, in alternate sections of heterozygous M50BL bulb, the immunostained glomeruli were the same as those labelled by *in situ* hybridization with the M50 probe (Fig. 2f, g, i, j). In each knock-in strain, one or two BL<sup>+</sup> glomeruli were seen at one medial and one lateral site in the bulb. These findings were consistent with our previous observations that BL expression in OmpBL mice does not alter OR gene expression in the nose or the targeting of axons to specific glomeruli in the bulb<sup>28</sup>.

In the glomeruli of the bulb, the axons of olfactory sensory neurons form synapses with the dendrites of bulb mitral and tufted relay neurons<sup>32</sup>. In M5BL and M50BL mice, BL was detected in approximately 35–40 mitral and tufted neurons near each BL<sup>+</sup> glomerulus. This is similar to the number of these neurons that innervate individual rat glomeruli<sup>33,34</sup>. The BL<sup>+</sup> neurons were invariably located near labelled glomeruli (Fig. 2k, l). In contrast, in OmpBL mice, BL<sup>+</sup> mitral and tufted neurons are seen throughout the bulb<sup>28</sup>. The patterns of labelling observed in the knock-in mice indicate that BL was transferred from olfactory sensory neurons expressing M5 or M50 ORs to the mitral and tufted relay neurons with which they form synapses in the olfactory bulb.

### OR inputs are clustered in the olfactory cortex

Mitral and tufted relay neurons in the bulb transmit signals to the olfactory cortex. The olfactory cortex is composed of several anatomically distinct areas, including the piriform cortex (the largest area), olfactory tubercle, anterior olfactory nucleus, and specific parts of the amygdala and entorhinal cortex<sup>6–9</sup>. The anterior and posterior halves of the piriform cortex differ morphologically and may be functionally distinct. Mitral cells project axons to the entire olfactory cortex, but tufted cells project only to the most anterior areas, such as the anterior olfactory nucleus and olfactory tubercle.

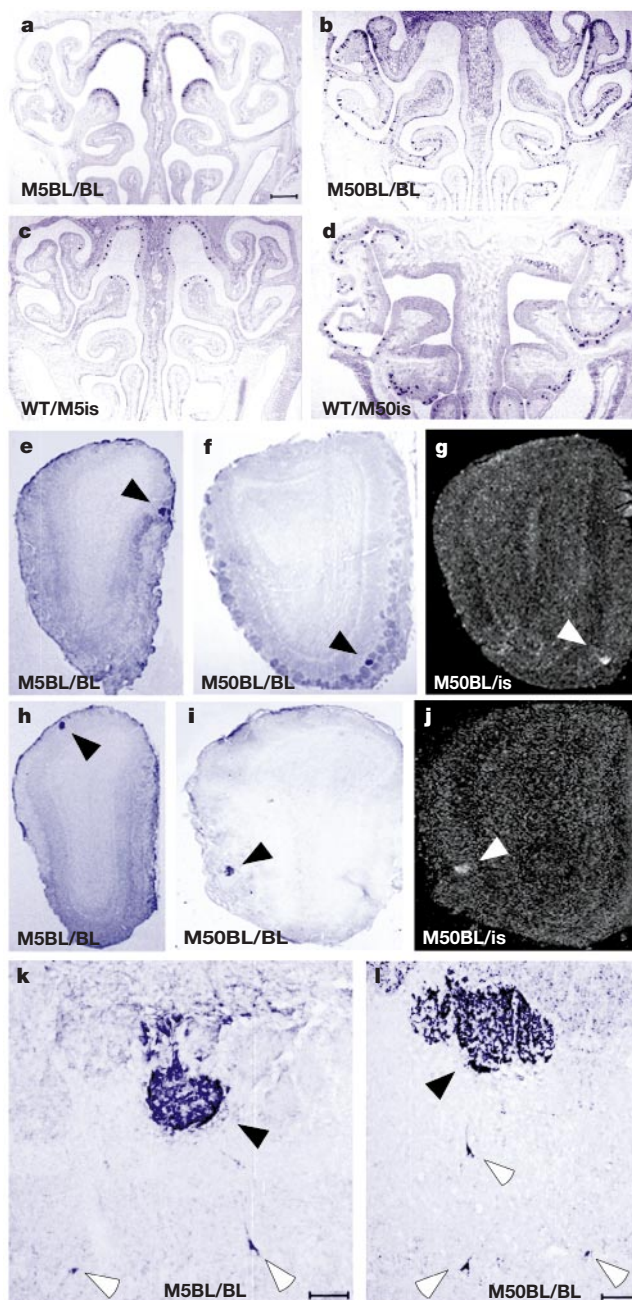
In the piriform cortex, mitral cell axons travel in the lateral



**Figure 1** Gene targeting strategy used to coexpress barley lectin (BL) with the *M5* or *M50* odorant receptor gene. Using homologous recombination in embryonic stem cells, an IRES/BL/loxP/tk-neo/loxP sequence (**b, f**) was inserted immediately 3' to the *M5* or *M50* odorant receptor coding region (**a, c, e, g**). The tk-neo cassette was removed by transient transfection with a Cre recombinase expression plasmid, leaving a single loxP site (**d, h**). The *M5* or *M50* coding region is represented by a black box; the IRES sequence by a red box; the BL coding region by a blue box; the tk-neo cassette by a grey box; and loxP sites by triangles. Thick black horizontal bars show regions used as probes in Southern blot screening for homologous recombinants. Restriction enzyme sites: A, *Afl*II; H, *Hind*III; K, *Kpn*I; N, *Not*I; X, *Xba*I; Xh, *Xho*I.

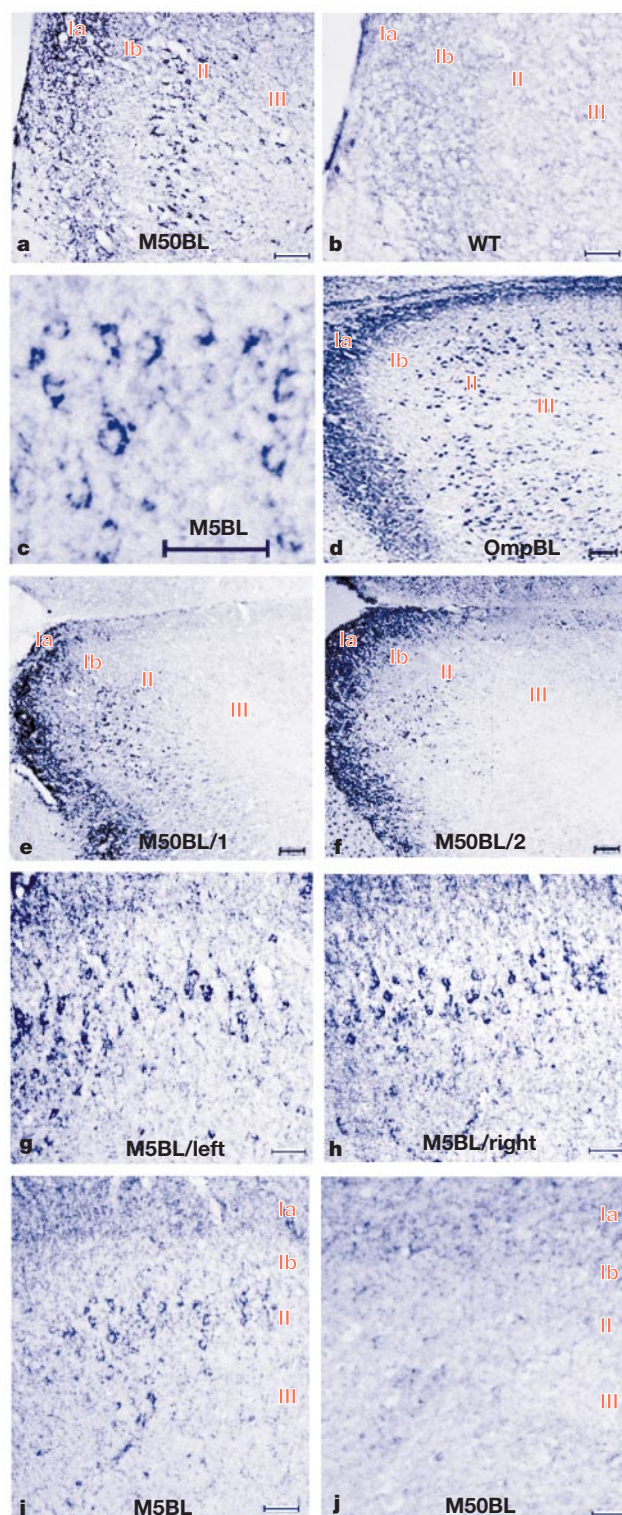
olfactory tract and then fan out over the cortical surface, branching to form synapses with the dendrites of cortical pyramidal neurons<sup>6–9</sup>. The synapses are formed in cortical layer Ia, whereas pyramidal neuron cell bodies are located in layers II and III (refs 6–9). In OmpBL mice, BL-labelled cells are seen in layers II and III throughout the entire olfactory cortex, consistent with the transfer of BL from mitral cells to the pyramidal neurons with which they form synapses (Fig. 3d)<sup>28</sup>.

In M5BL and M50BL knock-in mice, BL-labelled neurons were



**Figure 2** Detection of BL in the nose and olfactory bulb. Patterns of BL labelling in the nose of M5BL (**a**) or M50BL (**b**) knock-in mice resembled those seen in wild-type (WT) mice by *in situ* hybridization (is) with an M5 (**c**) or M50 (**d**) cRNA probe. A few glomeruli (black arrowheads) were BL-labelled at characteristic locations on the medial (**e, f**) and lateral (**h, i**) sides of the olfactory bulb in M5BL (**e, h**) and M50BL (**f, i**) mice. The M50 probe hybridized to the same glomeruli seen in **f** and **i** in adjacent sections (**g, j**), but not to other glomeruli. BL-labelled mitral and tufted cells (white arrowheads) were seen underlying the labelled glomeruli (black arrowheads) (**k, l**). Scale bars: **a**, 320  $\mu$ m; **k, l**, 40  $\mu$ m. Panels **a–j** are to the same scale.





**Figure 3** BL-labelled neurons in the anterior piriform cortex of M5BL and M50BL mice. BL<sup>+</sup> cells were arranged in clusters in layers II and III of the anterior piriform cortex of M5BL (**c, g, h, i**) and M50BL (**a, e, f**) mice. No BL-labelling was seen in wild-type (WT) mice (**b**), whereas in OmpBL mice, BL<sup>+</sup> cells were distributed throughout the anterior piriform cortex in layers II and III (**d**). The BL-labelled clusters had the same locations in different individuals of the same knock-in strain (**e, f**) and most were bilaterally symmetrical in the left and right hemisphere of the brain (**g, h**). The locations of clusters differed in the two strains (**i** (M5BL) compared with **j** (M50BL)). Scale bars: **a–c, g–j**, 40  $\mu$ m; **d–f**, 80  $\mu$ m.

also seen in layers II and III of the piriform cortex and other olfactory cortical areas (Figs 3–5). However, there were far fewer BL<sup>+</sup> cells in the knock-in mice than in OmpBL mice (Fig. 3d–f; see also below). Labelled cells were not observed in wild-type animals (Fig. 3b) nor were any cells expressing M5 or M50 detected by *in situ* hybridization in the brains of the knock-in animals (data not shown).

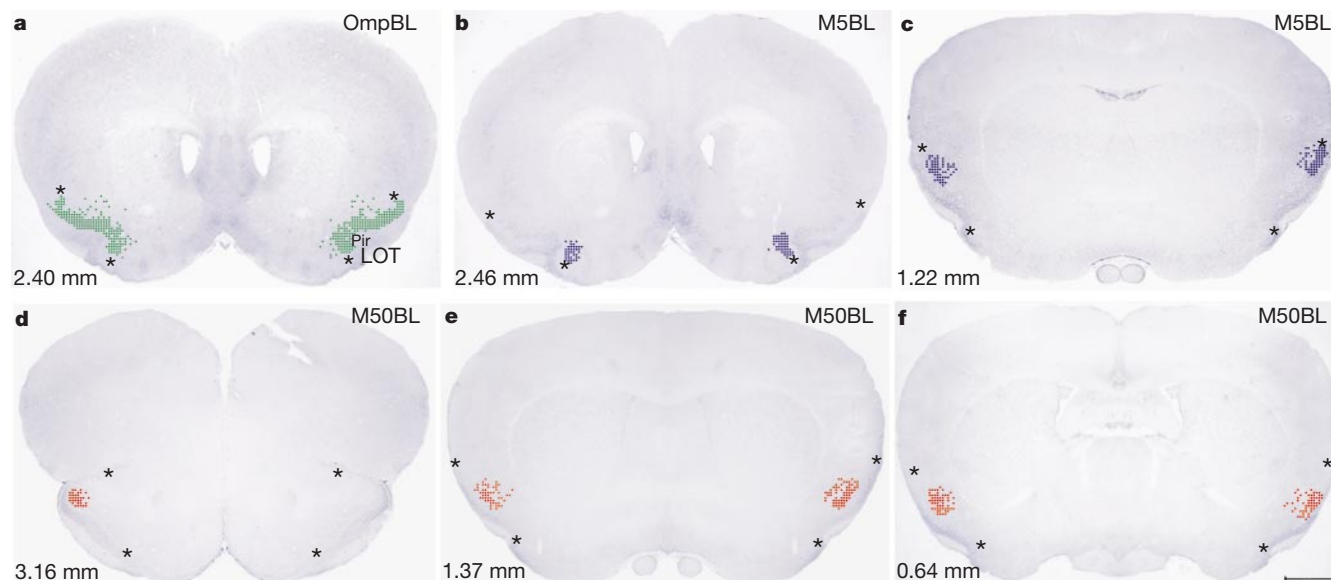
Notably, in the knock-in mice, the BL<sup>+</sup> neurons were organized in discrete clusters in most olfactory cortical areas. In M5BL mice, there were two clusters of BL<sup>+</sup> neurons in the anterior piriform cortex (Figs 3g–i and 4b, c), one in the olfactory tubercle (Fig. 5a), one in the anterior olfactory nucleus and one in the entorhinal cortex (Fig. 5e). M50BL animals had three labelled clusters in the anterior piriform cortex (Figs 3a, e, f and 4d–f), one in the olfactory tubercle (Fig. 5b) and two in the entorhinal cortex (Fig. 5f). In the posterior piriform cortex of both strains, there were numerous labelled neurons. Although there appeared to be a cluster of more densely labelled neurons here in each strain, there were also many less densely labelled neurons scattered over much of this area (Fig. 5c, d). The patterns of labelled neurons seen in olfactory cortex are depicted in Fig. 6. In both strains, a few labelled neurons were also detected in the horizontal limb of the diagonal band, which innervates the bulb. In addition, weakly labelled cells were sometimes seen in the neocortex in M5BL animals and, in M50BL mice, BL<sup>+</sup> neurons were detected in the paraventricular and arcuate nuclei of the hypothalamus (data not shown). Given the lack of evidence for a direct bulb–hypothalamus projection<sup>35</sup>, this might reflect a polysynaptic transport of the tracer.

Although some olfactory cortical neurons project axons to the bulb<sup>6–9,32</sup>, it is unlikely that BL was retrogradely transported to the cortex through these axons, as they terminate deep in the bulb, distant from BL<sup>+</sup> glomeruli and mitral/tufted neurons<sup>32</sup>. Moreover, the amygdala, which sends axons to the bulb, lacked BL<sup>+</sup> cells in the knock-in mice. In OmpBL mice, BL that was retrogradely transported from the bulb to neuromodulatory brain centres appeared to be relayed to many brain structures<sup>28</sup>, but no evidence of this was seen in the knock-in mice, perhaps because there was much less BL in the bulb.

### The olfactory cortex has a stereotyped sensory map

We next compared the locations of the BL<sup>+</sup> neuronal clusters in different animals ( $n = 5$  per strain). We focused first on the piriform cortex, the principal olfactory cortical area. Locations along the anterior–posterior axis were calculated from the number of tissue sections between the centre of a cluster and the end of the lateral olfactory tract (eLOT), a landmark that divides the anterior and posterior piriform cortex (Table 1). An atlas of the mouse brain was used to further examine the anterior–posterior locations of the clusters and to determine their positions along the dorsal–ventral and medial–lateral axes of individual olfactory cortical areas<sup>36</sup>.

These comparisons revealed an unexpected level of organization in the olfactory cortex. Detailed analyses of the clusters seen in the anterior piriform cortex indicated that the BL<sup>+</sup> neuronal clusters in this area had nearly identical locations in individuals of the same knock-in strain (Fig. 3e, f; see also Table 1). In addition, with the exception of one cluster that was invariably seen in only the left hemisphere of the brain of M50BL mice, the clusters had bilaterally symmetrical locations in the left and right hemispheres (Figs 3g, h and 4; see also Table 1). For example, in one M5BL mouse, the centre of the first (most anterior) M5BL cluster was 2.45 and 2.48 mm from eLOT in the left and right hemisphere, respectively, and 2.40 and 2.42 mm from this landmark in the left and right hemispheres of a different M5BL mouse (Table 1). The number of labelled neurons in each cluster was also similar among individuals and in the two hemispheres. For example, the numbers of BL<sup>+</sup> neurons in the first cluster in the left and right hemispheres of two M5BL mice were 1,926, 1,800, 1,566 and 1,776.



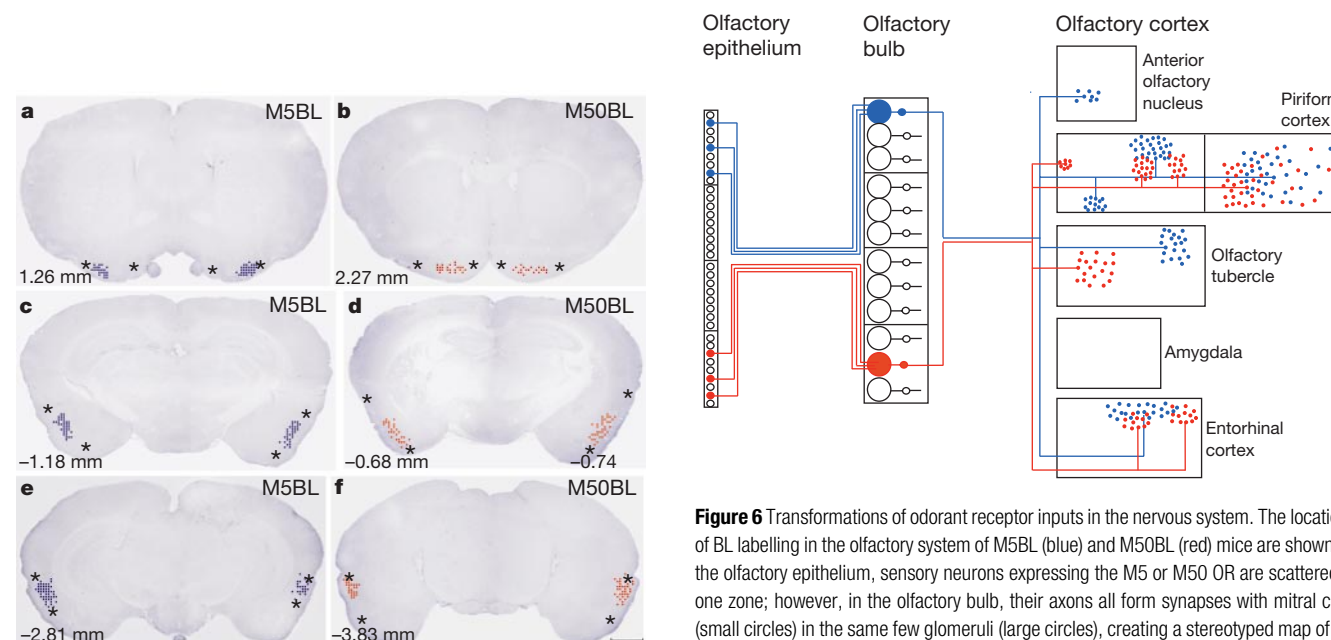
**Figure 4** Locations of labelled neurons in the anterior piriform cortex. The locations of BL-labelled neurons in high-magnification photographs of brain sections were marked by coloured dots, scaled, and then transferred to low-magnification photographs of the same sections. **a–f**, Coronal sections from M5BL (**a**), M5BL (**b, c**) and M50BL (**d–f**) mice. The

dorsal surface is at the top. The anterior–posterior distance of each section from eLOT is shown in millimetres. The dorsal and ventral limits of the piriform cortex (Pir) are indicated by asterisks. LOT, lateral olfactory tract. Scale bar, 640  $\mu$ m.

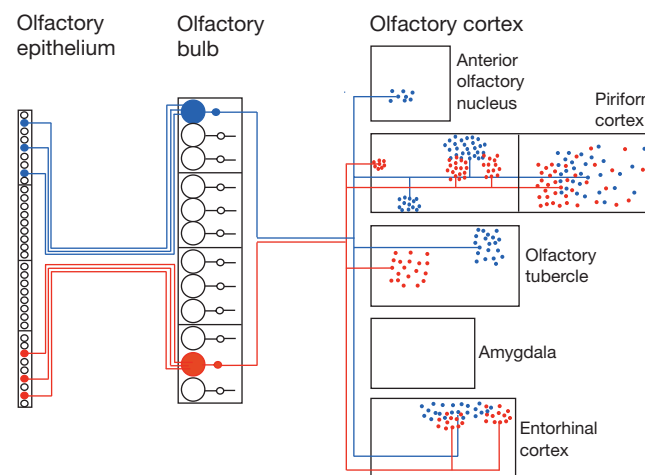
The positions of labelled neuronal clusters in the posterior piriform cortex, olfactory tubercle, anterior olfactory nucleus and entorhinal cortex were also similar in different individuals, and most were bilaterally symmetrical in the two hemispheres (Fig. 5). The only exception was in the entorhinal cortex, where labelled clusters appeared to have slightly different locations in the left and right hemispheres. Notably, the centres of the BL<sup>+</sup> clusters had different locations in the two knock-in strains. For example, in the anterior piriform cortex, the centres of clusters in M5BL mice were

$2.44 \pm 0.03$  and  $1.16 \pm 0.06$  mm from eLOT, whereas, in M50BL mice, they were  $3.15 \pm 0.06$ ,  $1.34 \pm 0.07$  and  $0.59 \pm 0.06$  mm from eLOT (Table 1). These clusters were also centred at different locations along the dorsal–ventral and medial–lateral axes in the two strains (Figs 4–6).

These results indicate that there is a stereotyped map of sensory inputs in the olfactory cortex. In this map, inputs from individual



**Figure 5** BL-labelled neurons in other olfactory cortical areas. **a–f**, The locations of labelled neurons, determined as in Fig. 4, are shown as coloured dots in photographs of coronal sections through the olfactory tubercle (**a, b**), posterior piriform cortex (**c, d**) and entorhinal cortex (**e, f**) of M5BL (**a, c, e**) and M50BL (**b, d, f**) mice. The dorsal–ventral or medial–lateral limits of each of these areas are indicated by asterisks. The dorsal surface is at the top. Scale bar, 640  $\mu$ m.



**Figure 6** Transformations of odorant receptor inputs in the nervous system. The locations of BL labelling in the olfactory system of M5BL (blue) and M50BL (red) mice are shown. In the olfactory epithelium, sensory neurons expressing the M5 or M50 OR are scattered in one zone; however, in the olfactory bulb, their axons all form synapses with mitral cells (small circles) in the same few glomeruli (large circles), creating a stereotyped map of OR inputs. In the nose and bulb, inputs from different ORs are segregated. Mitral cells carrying input from M5 or M50 form synapses with clusters of neurons (coloured dots) at specific sites in multiple olfactory cortical areas (boxes), creating another stereotyped map of sensory information. In this map, inputs from different ORs overlap spatially and single neurons may receive combinatorial input from multiple ORs. Inputs from the same ORs reach different olfactory cortical areas, allowing for parallel and perhaps differential processing of the same sensory inputs.



ORs are targeted to multiple, discrete clusters of cortical neurons, whose locations are similar or identical in different individuals.

### Combinatorial OR inputs in the olfactory cortex

While inputs from different ORs are spatially segregated in different glomeruli in the olfactory bulb, they appear to overlap, at least partially, in the olfactory cortex. Although the centres of labelled neuronal clusters differed in the two knock-in strains, partial overlap seems probable in many cases. For example, in the anterior piriform cortex, the centres of the second clusters in M5BL and M50BL mice were  $1.16 \pm 0.06$  and  $1.34 \pm 0.07$  mm from eLOT, respectively (Table 1). However, the clusters appeared to be partially overlapping in the dorsal–ventral dimension (Figs 4 and 6) and the M5BL cluster extended for  $0.86 \pm 0.02$  mm and the M50BL cluster for  $0.39 \pm 0.06$  mm (Table 1) along the anterior–posterior axis, suggesting that the two partially overlap along this axis as well (Fig. 6). Evidence for partially overlapping clusters in the two strains was also seen in the posterior piriform cortex and the entorhinal cortex, but not in the olfactory tubercle or the most anterior piriform cortex (Figs 5 and 6).

The proportion of the anterior piriform cortex occupied by neuronal clusters in the knock-in strains is consistent with overlapping inputs from different ORs. If this region received input from all of the approximately 1,000 ORs, and those inputs were spatially segregated, the labelled clusters could theoretically occupy approximately 0.1% of its total area of about  $10.5 \text{ mm}^2$  (see Methods)—about  $0.01 \text{ mm}^2$ . However, together the labelled clusters in this region in M5BL and M50BL mice occupied about 0.59 (5.7%) and  $0.44 \text{ mm}^2$  (4.2%), respectively. Similarly, in the olfactory tubercle, labelled clusters in the knock-in strains occupied roughly 1.7% (M5BL) and 7.5% (M50BL) of the total area.

Do individual cortical neurons receive inputs from multiple ORs? We determined that there were  $179,570 \pm 3,935$  labelled neurons in

the anterior piriform cortex of OmpBL mice ( $n = 2$ ), in which BL is expressed in all olfactory sensory neurons. In the same area, there were  $6,570 \pm 217$  labelled neurons in M5BL mice ( $n = 2$ ) and  $4,390 \pm 179$  in M50BL animals ( $n = 2$ ). This is markedly different from the number of BL<sup>+</sup> neurons that one might expect to see in knock-in mice (about 180) if each cortical neuron received input from only one of 1,000 ORs. Although BL released from mitral cell axon terminals in layer Ia could conceivably be taken up by the intermingled dendrites of both postsynaptic and non-postsynaptic cortical neurons, a more probable explanation is that individual cortical neurons receive input from a combination of different ORs. Further studies will be needed to clarify this point.

### Do all cortical areas receive inputs from the same ORs?

Does each olfactory cortical area receive information from the entire OR repertoire, or from only a subset of ORs? The patterns of labelled neurons seen in the knock-in animals clearly indicate that input from a single OR can be routed to multiple different areas of the olfactory cortex (Fig. 6). Both M5BL and M50BL mice had clusters of BL<sup>+</sup> neurons in the piriform cortex, olfactory tubercle and entorhinal cortex. However, the anterior olfactory nucleus contained a small cluster of labelled neurons in M5BL but not M50BL animals. In addition, the amygdala was devoid of BL<sup>+</sup> neurons in both knock-in strains, even though numerous labelled neurons are seen in amygdaloid nuclei that receive input from the main olfactory bulb in OmpBL mice<sup>28,29</sup>. It may be that there were BL-containing neurons in these areas that were not detected, because they contained only a small amount of tracer and/or were too few in number or too scattered to be noticed. However, another possibility that cannot be excluded is that some (or all) olfactory cortical areas receive input from only a subset of ORs.

### Discussion

Here, we used a genetic tracing system to investigate how sensory information is organized in the olfactory cortex. In knock-in mice that coexpressed a transneuronal tracer (BL) with one of about 1,000 different odorant receptor genes, the tracer travelled from sensory neurons in the nose that express the OR gene to second-order neurons in the olfactory bulb, and from the bulb neurons to third-order neurons in the olfactory cortex. This allowed visualization of neurons in the olfactory cortex that receive input from a particular OR.

These studies indicate that there is a stereotyped sensory map in the olfactory cortex. In this map, input from one OR is mapped onto clusters of neurons at a limited number of sites in the olfactory cortex. Sites that receive input from a given OR have similar or identical locations in different individuals, and most are bilaterally symmetrical in the two hemispheres of the brain. The existence of a stereotyped map in the cortex suggests a mechanism by which odorants could elicit similar perceptions, and perhaps emotional and physiological responses, in different individuals.

In marked contrast to the olfactory bulb, where inputs from different ORs are segregated in different glomeruli and relay neurons<sup>3,13–15,22</sup>, it seems that inputs from different ORs are mapped onto partially overlapping clusters of neurons in the olfactory cortex. In addition, it is possible that individual cortical neurons receive input from many different ORs. Given that different odorants are recognized by different combinations of ORs<sup>11</sup>, such an arrangement could permit an integration of the components of each odorant's unique receptor code that is important to the generation of diverse odour perceptions.

These studies also showed that input from one OR can be routed to multiple olfactory cortical areas as well as to several sites within one area. Considering that different olfactory cortical areas can transmit information to different target regions in the brain, this presumably allows input from the same OR to ultimately reach multiple brain regions that may serve different functions. Indeed,

**Table 1** Locations of BL<sup>+</sup> clusters in the anterior piriform cortex

| M5BL      | Distance to eLOT (μm)* | Length (μm)† | M50BL     | Distance to eLOT (μm)* | Length (μm)† |
|-----------|------------------------|--------------|-----------|------------------------|--------------|
| Cluster 1 |                        |              | Cluster 1 |                        |              |
| 1-L       | 2,450                  | 630          | 1-L       | 3,178                  | 420          |
| 1-R       | 2,478                  | 628          | 2-L       | 3,136                  | 490          |
| 2-L       | 2,401                  | 462          | 3-L       | 3,208                  | 336          |
| 2-R       | 2,415                  | 465          | 4-L       | 3,045                  | 385          |
| 3-L       | 2,485                  | 455          | 5-L       | 3,180                  | 350          |
| 3-R       | 2,485                  | 455          | Average   | 3,149                  | 396          |
| 4-L       | 2,415                  | 455          | ±s.d.     | 64                     | 62           |
| 4-R       | 2,413                  | 449          | Cluster 2 |                        |              |
| 5-L       | 2,457                  | 458          | 1-L       | 1,351                  | 432          |
| 5-R       | 2,415                  | 455          | 1-R       | 1,379                  | 490          |
| Average   | 2,441                  | 491          | 2-L       | 1,281                  | 420          |
| ±s.d.     | 33                     | 73           | 2-R       | 1,295                  | 478          |
| Cluster 2 |                        |              | 3-L       | 1,463                  | 343          |
| 1-L       | 1,211                  | 868          | 3-R       | 1,428                  | 350          |
| 1-R       | 1,225                  | 860          | 4-L       | 1,260                  | 350          |
| 2-L       | 1,085                  | 826          | 4-R       | 1,251                  | 329          |
| 2-R       | 1,099                  | 834          | 5-L       | 1,374                  | 336          |
| 3-L       | 1,085                  | 867          | 5-R       | 1,325                  | 385          |
| 3-R       | 1,090                  | 871          | Average   | 1,341                  | 391          |
| 4-L       | 1,187                  | 865          | ±s.d.     | 71                     | 60           |
| 4-R       | 1,192                  | 868          | Cluster 3 |                        |              |
| 5-L       | 1,205                  | 871          | 1-L       | 637                    | 283          |
| 5-R       | 1,213                  | 874          | 1-R       | 651                    | 319          |
| Average   | 1,159                  | 860          | 2-L       | 539                    | 280          |
| ±s.d.     | 61                     | 17           | 2-R       | 553                    | 322          |
|           |                        |              | 3-L       | 638                    | 210          |
|           |                        |              | 3-R       | 648                    | 210          |
|           |                        |              | 4-L       | 490                    | 210          |
|           |                        |              | 4-R       | 516                    | 224          |
|           |                        |              | 5-L       | 620                    | 280          |
|           |                        |              | 5-R       | 613                    | 315          |
|           |                        |              | Average   | 591                    | 265          |
|           |                        |              | ±s.d.     | 60                     | 47           |

Data obtained from the left (L) and right (R) hemispheres of the brains of five mice (3 female and 2 male) of each knock-in strain. Clusters are numbered from anterior to posterior.

\* Distance from the centre of each cluster of BL<sup>+</sup> neurons to the posterior end of the lateral olfactory tract (eLOT).

† Anterior–posterior length of each cluster.



although information from most olfactory cortical areas is relayed to the frontal cortex and lateral hypothalamus<sup>6–9</sup>, the anterior and posterior piriform cortex transmit signals to different frontal cortex regions<sup>37</sup>, and only the olfactory tubercle relays information to the nuclei gemini of the hypothalamus<sup>38</sup>. Divergent inputs to different olfactory cortical areas should, in addition, permit a parallel—and possibly differential—processing of OR inputs in which those inputs may be combined and modulated in different ways for transmission to either the same or different areas of the neocortex and limbic system.

It cannot be excluded, however, that some (or all) cortical areas receive input from only a subset of ORs. This is one possible explanation for the absence of labelled neurons in the amygdala in the knock-in strains. Given the importance of the amygdala in emotional states<sup>1,2,39</sup> and its projection to parts of the hypothalamus that do not receive input from other olfactory cortical areas, it may be that this area receives input only from ORs that convey information relevant to particular emotional states or to instinctive behaviours or physiological states that are governed by the hypothalamus<sup>1,2,40,41</sup>.

The generation of diverse olfactory perceptions appears to arise from a series of transformations of sensory inputs that involve the convergence, divergence and parallel processing of signals derived from the OR family (Fig. 6). In the nose, thousands of neurons that receive input from the same OR are scattered in one zone. In the olfactory bulb the axons of those neurons converge in a few specific glomeruli to form synapses with 50–100 relay neurons, each dedicated to input from one OR. In the olfactory cortex, inputs from a particular OR diverge to thousands of neurons located in stereotypical clusters in multiple olfactory cortical areas. Here, inputs from different ORs overlap spatially and might also be combined in single cortical neurons. Information from the same OR is processed in parallel in different olfactory cortical areas, potentially allowing OR inputs to be integrated or refined in different ways before delivery to the neocortex and limbic system. □

## Methods

### Generation of M5BL and M50BL mice

Clones containing *M5* or *M50* odorant receptor genes were isolated from a mouse (129/SvJ) genomic library (Stratagene) screened with a *M5* or *M50* cDNA probe<sup>10,31</sup>. An *AscI* restriction enzyme site was created just 3' to the stop codon of the *M5* or *M50* gene using PCR<sup>42</sup>. We used the plasmid pETLpA/LTNL<sup>15</sup> to generate an IRES/BL/loxP/tk-neo/loxP (IBL/LTNL) construct. Briefly, the tau-lacZ portion of pETLpA/LTNL was replaced with a truncated BL cDNA<sup>28</sup>, generating plasmid pIBL/LTNL. The IBL/LTNL fragment was excised from the plasmid with *AscI* and inserted into the unique *AscI* site 3' to the *M5* or *M50* coding region to give the targeting vector pM5BL or pM50BL. Embryonic stem cells were electroporated with the constructs, cultured with G418, and individual clones were isolated and analysed by Southern blotting to identify homologous recombinants<sup>43</sup>. To remove the floxed tk-neo cassette, cells were electroporated with a Cre expression plasmid (GIBCO), cultured with Gancyclovir (Syntex Chemicals), and then analysed by Southern blotting. Transgenic mouse lines (M5BL and M50BL) were generated<sup>43</sup> to give mice in a mixed (129/SvJ) × C57BL/6 background.

### Immunohistochemistry

Mice were perfused with 4% paraformaldehyde, the brain and nose were isolated, fixed in 4% paraformaldehyde for an additional 2 h, soaked in 30% sucrose for 48 h, and then frozen in OCT (Sakura)<sup>10</sup>. Coronal cryostat sections (14 µm) were treated with goat anti-wheat germ agglutinin (WGA) antibodies (0.8 µg ml<sup>-1</sup>) (Vector Laboratories) (at room temperature for 2 h) and then with biotinylated anti-goat IgG (0.8 µg ml<sup>-1</sup>; Vector Laboratories) for 1 h (ref. 44). Sections were then treated according to manufacturers' instructions with components of the TSA (NEN Life Science), ABC and DAB (Vector Laboratories) kits, and examined microscopically.

### In situ hybridization

Fresh frozen sections (14 µm) were hybridized to fluorescein-labelled cRNA probes at 58 °C as described<sup>45</sup>. The *M5* probe was a 926-nucleotide coding region fragment whereas the *M50* probe included the coding region plus 700 nucleotides of the 5' and 300 nucleotides of the 3' untranslated region. After hybridization, sections were incubated with peroxidase-anti-fluorescein antibody (Roche; 1:750 dilution) at room temperature for 1 h, followed by TSA, ABC and DAB treatment as described above. *In situ* hybridization

with <sup>33</sup>P-labelled cRNA probes was performed on 10 µm paraffin sections of olfactory bulb from wild-type 129/SvJ and heterozygous knock-in mice as described<sup>10,13</sup>.

### Data analyses

Each brain was sectioned from the anterior tip of the olfactory bulb to the brain stem, and whenever possible all the sections were collected. Positions of a few missing sections were carefully recorded to accurately estimate anterior–posterior distances from the section numbers. Initially, two out of every five consecutive sections of each knock-in mouse brain were stained. If positive cells were detected, all sections around the positive cells were then immunostained. Knock-in mice from one week to four months of age were analysed. Some adult mice were caged with one member of the opposite sex, and some were exposed periodically to a mixture of odorants. Wild-type 129/SvJ mice, wild-type littermates of the knock-in mice and OmpBL transgenic mice of the same age and sex were treated the same way as the knock-in mice, and were used as controls throughout the studies. One out of every twenty sections was analysed in OmpBL mice. Counting of labelled cells was carried out blind, at least three times on different days.

We identified brain structures microscopically by reference to a mouse brain atlas<sup>36</sup>. The anterior–posterior lengths of brain structures and BL-labelled clusters were calculated from the number of sections multiplied by 14 µm. Dorsal–ventral and medial–lateral distances were measured on computer images of sections photographed with a size bar. To obtain the total area of the anterior piriform cortex or olfactory tubercle, we first used the brain atlas to measure the dorsal–ventral and medial–lateral extent of the surface of the structure at different points along the anterior–posterior axis and then determined the average. We then compared the anterior–posterior length of the structure in the atlas with that determined in the knock-in mice, and normalized the dorsal–ventral/medial–lateral extent of the structure accordingly.

Received 22 January; accepted 3 September 2001.

- Kandel, E., Schwartz, J. & Jessell, T. *Principles of Neural Science* 625–647 (McGraw-Hill, New York, 2000).
- Zigmond, M., Bloom, F., Landis, S., Roberts, J. & Squire, L. *Fundamental Neural Science* 737–753 (Academic, San Diego, 1999).
- Buck, L. B. The molecular architecture of odor and pheromone sensing in mammals. *Cell* **100**, 611–618 (2000).
- Buck, L. & Axel, R. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* **65**, 175–187 (1991).
- Mombaerts, P. Molecular biology of odorant receptors in vertebrates. *Annu. Rev. Neurosci.* **22**, 487–509 (1999).
- Scott, J. W. The olfactory bulb and central pathways. *Experientia* **42**, 223–232 (1986).
- Price, J. L. In *Neurobiology of Taste and Smell* (eds Finger, T. E. & Sliver, W. L.) 179–203 (Wiley, New York, 1987).
- Shipley, M. T. & Ennis, M. Functional organization of olfactory system. *J. Neurobiol.* **30**, 123–176 (1996).
- Haberly, L. B. In *The Synaptic Organization of the Brain* (ed. Shepherd, G. M.) 377–416 (Oxford Univ. Press, New York, 1998).
- Ressler, K. J., Sullivan, S. L. & Buck, L. B. A zonal organization of odorant receptor gene expression in the olfactory epithelium. *Cell* **73**, 597–609 (1993).
- Malnic, B., Hirono, J., Sato, T. & Buck, L. B. Combinatorial receptor codes for odors. *Cell* **96**, 713–723 (1999).
- Vassar, R., Ngai, J. & Axel, R. Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. *Cell* **74**, 309–318 (1993).
- Ressler, K. J., Sullivan, S. L. & Buck, L. B. Information coding in the olfactory system: evidence for a stereotyped and highly organized epitope map in the olfactory bulb. *Cell* **79**, 1245–1255 (1994).
- Vassar, R. *et al.* Topographic organization of sensory projections to the olfactory bulb. *Cell* **79**, 981–991 (1994).
- Mombaerts, P. *et al.* Visualizing an olfactory sensory map. *Cell* **87**, 675–686 (1996).
- Kauer, J. S., Senseman, D. M. & Cohen, L. B. Odor-elicited activity monitored simultaneously from 124 regions of the salamander olfactory bulb using a voltage-sensitive dye. *Brain Res.* **418**, 255–261 (1987).
- Katoh, K., Koshimoto, H., Tani, A. & Mori, K. Coding of odor molecules by mitral/tufted cells in rabbit olfactory bulb. II. Aromatic compounds. *J. Neurophysiol.* **70**, 2161–2175 (1993).
- Rubin, B. D. & Katz, L. C. Optical imaging of odorant representations in the mammalian olfactory bulb. *Neuron* **23**, 499–511 (1999).
- Uchida, N., Takahashi, Y. K., Tanifuji, M. & Mori, K. Odor maps in the mammalian olfactory bulb: domain organization and odorant structural features. *Nature Neurosci.* **3**, 1035–1043 (2000).
- Buck, L. B. Information coding in the vertebrate olfactory system. *Annu. Rev. Neurosci.* **19**, 517–544 (1996).
- Hildebrand, J. G. & Shepherd, G. M. Mechanisms of olfactory discrimination: converging evidence for common principles across phyla. *Annu. Rev. Neurosci.* **20**, 595–631 (1997).
- Mori, K., Nagao, H. & Yoshihara, Y. The olfactory bulb: coding and processing of odor molecule information. *Science* **286**, 711–715 (1999).
- Luskin, M. B. & Price, J. L. The distribution of axon collaterals from the olfactory bulb and the nucleus of the horizontal limb of the diagonal band to the olfactory cortex, demonstrated by double retrograde labeling techniques. *J. Comp. Neurol.* **209**, 249–263 (1982).
- Haberly, L. B. & Price, J. L. The axonal projection patterns of the mitral and tufted cells of the olfactory bulb in the rat. *Brain Res.* **129**, 152–157 (1977).
- Scott, J. W., McBride, R. L. & Schneider, S. P. The organization of projections from the olfactory bulb to the piriform cortex and olfactory tubercle in the rat. *J. Comp. Neurol.* **194**, 519–534 (1980).
- Ojima, H., Mori, K. & Kishi, K. The trajectory of mitral cell axons in the rabbit olfactory cortex revealed by intracellular HRP injection. *J. Comp. Neurol.* **230**, 77–87 (1984).
- Buonviso, N., Revial, M. F. & Jourdan, F. The projections of mitral cells from small local regions of the bulb: an anterograde tracing study using PHA-L (Phaseolus vulgaris leucoagglutinin). *Eur. J. Neurosci.* **3**, 493–500 (1991).

28. Horowitz, L. F., Montmayeur, J. P., Echelard, Y. & Buck, L. B. A genetic approach to trace neural circuits. *Proc. Natl Acad. Sci. USA* **96**, 3194–3199 (1999).
29. Yoshihara, Y. *et al.* A genetic approach to visualization of multisynaptic neural pathways using plant lectin transgene. *Neuron* **22**, 33–41 (1999).
30. Wang, F., Nemes, A., Mendelsohn, M. & Axel, R. Odorant receptors govern the formation of a precise topographic map. *Cell* **93**, 47–60 (1998).
31. Sullivan, S. L., Adamson, M. C., Ressler, K. J., Kozak, C. A. & Buck, L. B. The chromosomal distribution of mouse odorant receptor genes. *Proc. Natl Acad. Sci. USA* **93**, 884–888 (1996).
32. Shepherd, G. M. & Greer, C. A. in *The Synaptic Organization of the Brain* (ed. Shepherd, G. M.) 159–203 (Oxford Univ. Press, New York, 1998).
33. Allison, A. C. The structure of the olfactory bulb and its relationship to the olfactory pathways in the rabbit and the rat. *J. Comp. Neurol.* **98**, 309–353 (1953).
34. Royet, J., Distel, H., Hudson, R. & Gervais, R. A re-estimation of the number of glomeruli and mitral cells in the olfactory bulb of rabbit. *Brain Res.* **788**, 35–42 (1998).
35. Risold, P. Y., Thompson, R. H. & Swanson, L. W. The structural organization of connections between hypothalamus and cerebral cortex. *Brain Res. Brain Res. Rev.* **24**, 197–254 (1997).
36. Franklin, K. & Paxinos, G. *The Mouse Brain in Stereotaxic Coordinates* (Academic, San Diego, 1997).
37. Datiche, F. & Cattarelli, M. Reciprocal and topographic connections between the piriform and prefrontal cortices in the rat: a tracing study using the B subunit of the cholera toxin. *Brain Res. Bull.* **41**, 391–398 (1996).
38. Price, J. L., Slotnick, B. M. & Revial, M. F. Olfactory projections to the hypothalamus. *J. Comp. Neurol.* **306**, 447–461 (1991).
39. McDonald, A. J. Cortical pathways to the mammalian amygdala. *Prog. Neurobiol.* **55**, 257–332 (1998).
40. Simerly, R. in *The Rat Nervous System* (ed. Paxinos, G.) 353–376 (Academic, San Diego, 1995).
41. Armstrong, W. in *The Rat Nervous System* (ed. Paxinos, G.) 377–390 (Academic, San Diego, 1995).
42. Ausubel, F. *et al.* *Current Protocols in Molecular Biology* 15.1.1–15.8.7 (Wiley, New York, 2000).
43. Joyner, A. L. *Gene Targeting—A Practical Approach* 107–146 (IRL press, Oxford, 1993).
44. Cote, S., Ribeiro-Da-Silva, A. & Cuello, A. in *Immunohistochemistry* (ed. Cuello, A.) 147–168 (Wiley, Chichester, 1993).
45. Matsunami, H. & Buck, L. B. A multigene family encoding a diverse array of putative pheromone receptors in mammals. *Cell* **90**, 775–784 (1997).
46. Zhao, H. *et al.* Functional expression of a mammalian odorant receptor. *Science* **279**, 237–242 (1998).
47. Krautwurst, D., Yau, K. W. & Reed, R. R. Identification of ligands for olfactory receptors by functional expression of a receptor library. *Cell* **95**, 917–926 (1998).
48. Araneda, R. C., Kini, A. D. & Firestein, S. The molecular receptive range of an odorant receptor. *Nature Neurosci.* **3**, 1248–1255 (2000).

## Acknowledgements

We are grateful to F. Alt for providing training in his laboratory in embryonic stem cell techniques. We would like to thank P. Mombaerts for the pETLpA/LTNL plasmid; Q. Ma for helpful suggestions; and D. Hubel and R. Brent for critically reading the manuscript. We would also like to thank C. Gao for technical assistance, and members of the Buck laboratory for comments and discussions. This work was supported by the Howard Hughes Medical Institute, grants from the National Institutes of Health and the Department of Defense (L.B.B.), a Howard Hughes Predoctoral Fellowship (L.F.H.), the Fyssen Foundation (J.-P.M.) and the Alice and Joseph Brook Fund (J.-P.M.).

Correspondence and requests for material should be addressed to L.B.B. (e-mail: lbuck@hms.harvard.edu).

# Correlated fast X-ray and optical variability in the black-hole candidate XTE J1118+480

G. Kanbach\*, C. Straubmeier\*, H. C. Spruit† & T. Belloni‡

\* Max-Planck-Institut für Extraterrestrische Physik; and † Max-Planck-Institut für Astrophysik, Postfach 1317, 85741 Garching, Germany

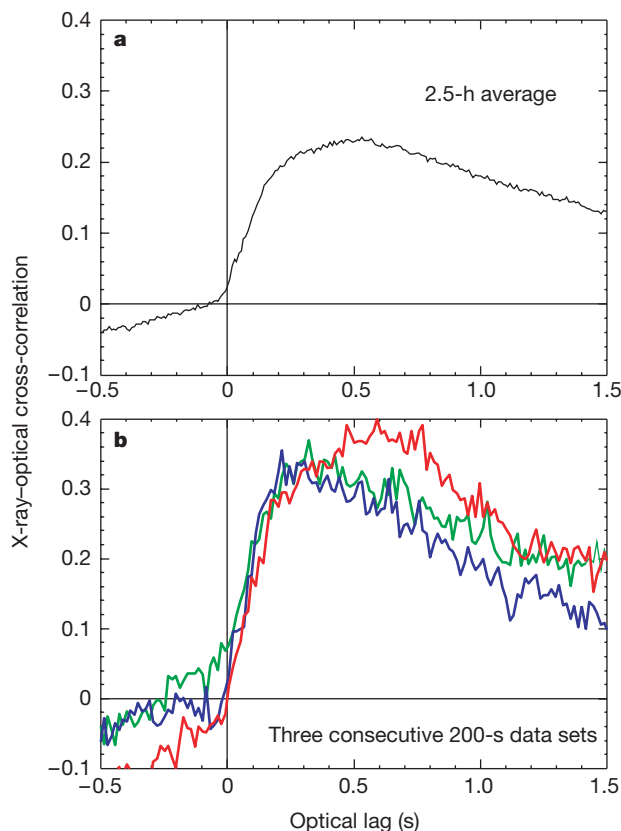
‡ Osservatorio Astronomico di Brera, via Bianchi 46, 23807 Merate (Lecco), Italy

Black holes become visible when they accrete gas, a common source of which is a close stellar companion. The standard theory for this process (invoking a ‘thin accretion disk’) does not explain some spectacular phenomena associated with these systems, such as their X-ray variability<sup>1</sup> and relativistic outflows<sup>2</sup>, indicating some lack of understanding of the actual physical conditions. Simultaneous observations at multiple wavelengths can provide strong constraints on these conditions. Here we report simultaneous high-time-resolution X-ray and optical observations of the transient source XTE J1118+480, which show a strong but puzzling correlation between the emissions. The optical emission rises suddenly following an increase in the X-ray output, but with a dip 2–5 seconds in advance of the X-rays. This result is not easy to understand within the simplest model of the optical emission, where the light comes from reprocessed X-rays. It is probably more consistent with an earlier suggestion<sup>3</sup> that the optical light is cyclosynchrotron emission that originates in a region about 20,000 km from the black hole. We propose that the time dependence is evidence for a relatively slow (<0.1c), magnetically controlled outflow.

The X-ray transient XTE J1118+480 (=KV UMa) provided a unique opportunity for simultaneous X-ray and optical observations because of its long duration (from January to August 2000), its proximity, and the lack of obscuration at its position above the galactic plane<sup>4–6</sup>. The orbit of the companion shows<sup>5</sup> that the central object has a mass  $>6M_{\odot}$ . We have obtained a total of 2.5 h of simultaneous observation distributed over four consecutive nights (4–7 July 2000). The X-rays were strongly variable, in a mode typical (see ref. 7, for example) for black-hole candidates (BHC) in the low-hard state (‘low’ intensity at  $\sim$  keV energies but ‘hard’ power-law spectrum extending to  $>10$  keV), like Cyg X-1. The amplitude of the optical variability was smaller, about 10% r.m.s. on timescales of minutes and shorter. Results of another simultaneous observation (with the Rossi X-ray Timing Explorer, RXTE, and the Hubble space telescope, HST) have already been announced<sup>8</sup>.

Correlations between the two time series are hard to identify reliably in the time series themselves, but are quite unambiguous by cross-correlation of segments as short as 1 min (Fig. 1). An optical response starts within about 30 ms following the X-rays, with a maximum reached after about 0.5 s. The rapid variability of the optical emission (Fig. 2) excludes intrinsic thermal emission from a cool accretion disk as the source of the (variable part of the) optical emission.

The shape of the cross-correlation as shown by Fig. 1 is quite suggestive of a simple explanation in terms of reprocessing<sup>9,10</sup>. X-rays for the central regions near the hole illuminate the upper layers of the surrounding accretion disk, and by heating these cause optical and ultraviolet radiation. The optical response is then easily modelled with an accretion disk inclined out of the plane of the sky. The rapid rise would be due to the relatively short light travel time delay of photons reprocessed on the Earth-facing half of the disk, and the slower decay arising from the longer path followed by photons reprocessed on the far side. On quantitative inspection,



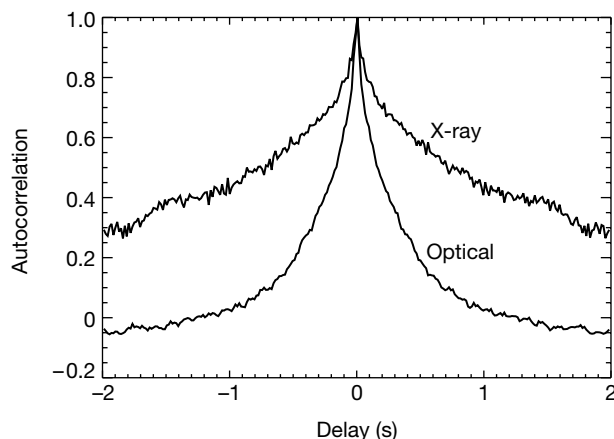
**Figure 1** Cross-correlation of the X-ray and optical time series of XTE J1118+480, showing onset of an optical response within 30 ms. Positive lag corresponds to delay of optical emission. **a**, Average of all 2.5 h of data. **b**, A selection of three consecutive 200-s segments of data, illustrating variability of the correlation. Piecewise linear fits to 30-s averages have been subtracted from both time series before correlating. The X-ray observations were made with the Rossi X-ray timing explorer (<http://heasarc.gsfc.nasa.gov/docs/xte/>), the optical observations with the Optima photometer<sup>21</sup> attached to the 1.3-m telescope at Mt Skinakas, Crete (<http://observatory.physics.uoc.gr/>). The photometer consists of a cluster of fibre-coupled avalanche photo diodes sensitive from 450 to 950 nm at a mean efficiency of 50%. Individual photon arrival times are recorded at 2  $\mu$ s absolute accuracy using a global positioning system (GPS)-based clock. Like in other observations of this source<sup>22</sup>, a weak quasiperiodic oscillation of 0.08–0.1 Hz was present in the X-rays.

however, this interpretation can be ruled out. First, the X-ray light curve does not have enough variability on timescales of 10–500 ms. With the observed X-ray autocorrelation (Fig. 2) a time lag of the order of a few seconds can be produced, but not the observed steep rise in the first few 100 ms. This is also evident from the optical variability itself. If the optical variation were due to reprocessing of the X-rays, its autocorrelation should be wider than that of the X-rays. This is the opposite of what is observed (Fig. 2).

Second, there is significant variability in the cross-correlation (Fig. 1), on timescales as short as 3 min. This is difficult to explain as a variation in the reprocessing properties of the disk, because at 0.2–2 lightseconds from the centre the intrinsic timescales of the disk are much longer than 3 min. Some variation could be due to changes in the X-ray autocorrelation function, but quantitatively this solution again suffers from the lack of sufficient variability of the X-rays on short timescales.

Third, in X-ray binaries where reprocessing is indicated by independent evidence (such as Bowen fluorescence<sup>11</sup>), the optical–X-ray flux ratio is only a few per cent. This is much less than is observed in XTE J1118+480, which had an optical luminosity<sup>12</sup> around  $10^{35}$  erg s<sup>–1</sup>, some 50% of the X-ray luminosity. Though



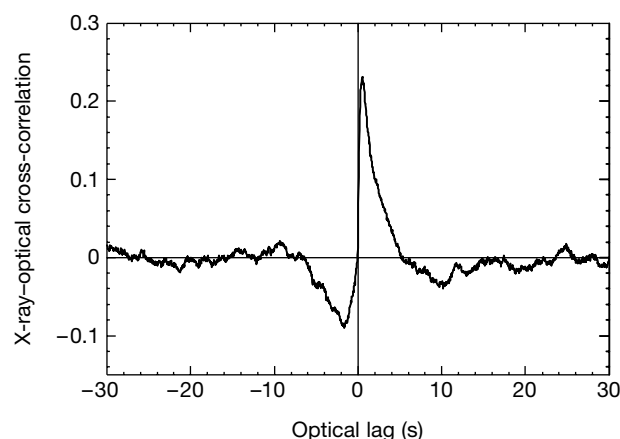


**Figure 2** Autocorrelations of the X-ray and optical time series, showing that the optical variability is not due to reprocessing of the X-rays. Timescales less than 70 ms are present in both the X-ray and optical light curves, but the optical autocorrelation is significantly narrower than the X-ray autocorrelation. If  $f_X(t)$ ,  $f_O(t)$  are the X-ray and optical light curves and  $g(t)$  is the optical response of the disk to an infinitesimally short spike of X-rays, then  $f_O = f_X * g$ , where  $*$  denotes convolution. The X-ray–optical cross-correlation is then  $C_{XO} = f_X * f_O = A_X * g$  and  $A_O = A_X * g * g$ , where  $A_{X,O} = f_{X,O} * f_{X,O}$  are the X-ray and optical autocorrelation functions. In a reprocessing model, the optical autocorrelation must therefore be broader than  $A_X$ .

our observations address only the (small) variable part of the optical emission, the unusually large optical flux itself is not easily explained either as reprocessed X-rays or as thermal emission from a cool disk.

A curious feature of the cross-correlation is the dip at 2–5 s before zero lag (Fig. 3). The optical emission appears to decrease before the onset of an X-ray increase. This has been seen before in the BHC GX 339-4 (the only other source for which simultaneous X-optical observations at millisecond time resolution have been reported<sup>13</sup>). This feature poses a challenge for any model of the optical emission. Variations in the optical emission preceding the X-rays suggest a source of variation further out in the disk where (some of the) optical light may be produced, for example by a change in the mass flow rate. However, the optical signal would then require a decreasing mass flux preceding an X-ray increase. Cooling of the X-ray-emitting region by inverse Compton scattering on optical and ultraviolet photons<sup>3</sup> would cause anticorrelated X-ray and optical variation, but any time delay would be much less than seconds.

Significant variability in the optical takes place on timescales as short as 100 ms (see Figs 1, 2); hence, the size of the optical emission region is not larger than a light travel distance of  $3 \times 10^9$  cm. The observed optical brightness  $\nu F_\nu = 1.5 \times 10^{-10}$  erg cm<sup>-2</sup> s<sup>-1</sup> then yields a minimum brightness temperature of  $2 \times 10^6$  K for the inferred distance of the source (1.8 kpc). The radiation mechanism that most plausibly produces such brightness is cyclotron emission in a strong magnetic field<sup>3</sup>. To explain the (variable) delay of the optical emission, we interpret the emission region as the photosphere (surface of optical depth unity) in a magnetically dominated outflow from the central regions of the accretion. Modulations in the accretion rate onto the hole, evident in the X-ray light curve, are assumed to propagate into the outflow as modulations of the flow velocity and mass flux. Steepened into shocks, these modulations produce the observed radiation, as in the standard internal shock model of jet emission<sup>14,15</sup>. The optical variability reflects the passage of these modulations through the photosphere, and the optical delay of about 0.5 s is identified with the travel time of the flow to the photosphere from its origin near the hole. We find that this model is realistic only if the outflow has a



**Figure 3** Correlation on timescales up to 30 s, showing the enigmatic ‘precognition dip’ at negative lags of 2–5 s. In the available models for the radiation from black hole candidates the optical light is produced either simultaneously with the X-rays, or later, by reprocessing of X-rays. Such models do not explain how the optical emission can ‘know something’ about X-ray emission that comes later.

rather low velocity,  $v < 30,000$  km s<sup>-1</sup>. If the assumed velocity is too large, for example mildly relativistic ( $\beta \approx 0.5$ ) the photosphere is at such a distance (about  $10^{10}$  cm) that the magnetic field of around  $10^6$  G which is required to produce enough optical emission would be unrealistically high.

The outflow proposed here is not the same as that producing the observed radio emission, which is likely to originate from a less dense but faster flow, possibly a mildly relativistic jet<sup>16,17</sup>. Both kinds of outflow may be present at the same time, concentric to each other. The flow inferred here would then be the microquasar analogue of, for example, the broad-line region outflows<sup>18,19</sup> of quasars, or the ‘anomalous outflows’<sup>20</sup> of SS433. The cyclotron outflow model is still tentative, however, as it does not naturally explain the prominent ‘precognition dip’ at negative lags. □

Received 3 July; accepted 5 October 2001.

- Lewin, W. H. G., van Paradijs, J. & van den Heuvel, E. P. J. (eds) *X-ray Binaries* (Cambridge Univ. Press, Cambridge, 1995).
- Mirabel, I. F. & Rodríguez, L. F. Sources of relativistic jets in the galaxy. *Annu. Rev. Astron. Astrophys.* **37**, 409–443 (1999).
- Fabian, A. C. et al. GX 339-4—Cyclotron radiation from an accretion flow. *Astron. Astrophys.* **111**, L9–L12 (1982).
- Remillard, R. A., Morgan, E., Smith, D. & Smith, E. XTE 1118+480. *IAU Circ.* 7389 (2000).
- McClintock, J. E. et al. A black hole of  $>6M_\odot$  in the X-ray nova XTE 1118+480. *Astrophys. J.* **551**, L147–L150 (2001).
- Wagner, R. M. et al. The halo black-hole X-ray transient XTE 1118+480. Preprint astro-ph/0104032 at (<http://xxx.lanl.gov>) (2001).
- Van der Klis, M. in *X-ray Binaries* (eds Lewin, W. H. G., van Paradijs, J. & van den Heuvel, E. P. J.) 252–300 (Cambridge Univ. Press, Cambridge, 1995).
- Haswell, C. A., Skillman, D., Patterson, J., Hynes, R. I. & Cui, W. XTE J1118+480. *IAU Circ.* 7427 (2000).
- O’Brien, K. & Horne, K. in *Ross 2000: Astrophysics with the Rossi X-ray Timing Explorer* E91–E93 (NASA Goddard Space Flight Center, Greenbelt, Maryland, 2000).
- Hynes, R. I., O’Brien, K., Horne, K., Chen, W. & Haswell, C. A. Echoes from an irradiated disc in GRO J1655-40. *Mon. Not. R. Astron. Soc.* **299**, L37–L41 (1998).
- Van Paradijs, J. E. & McClintock, J. in *X-ray Binaries* (eds Lewin, W. H. G., van Paradijs, J. & van den Heuvel, E. P. J.) 58–121 (Cambridge Univ. Press, Cambridge, 1995).
- McClintock, J. E. et al. Complete and simultaneous spectral observations of the black-hole X-ray nova XTE J1118+480. *Astrophys. J.* **555**, 477–482 (2001).
- Motch, C., Ilovaisky, S. A. & Chevalier, C. Discovery of fast optical activity in the X-ray source GX 339-4. *Astron. Astrophys.* **109**, L1–L4 (1982).
- Fender, R. P. et al. Spectral evidence for a powerful compact jet from XTE J1118+480. *Mon. Not. R. Astron. Soc.* **322**, L23–L27 (2001).
- Markoff, S., Falcke, H. & Fender, R. A jet model for the broadband spectrum of XTE J1118+480. Synchrotron emission from radio to X-rays in the Low/Hard spectral state. *Astron. Astrophys.* **372**, L25–L28 (2001).
- Rees, M. J. The M87 jet—Internal shocks in a plasma beam. *Mon. Not. R. Astron. Soc.* **184**, 61–65 (1978).
- Kaiser, C. R., Sunyaev, R. A. & Spruit, H. C. Internal shock model for microquasars. *Astron. Astrophys.* **356**, 975–988 (2000).

18. Leahy, J. P. in *Astrophysical Discs—An EC Summer School* (eds Sellwood, J. A. & Goodman, J.) 246–250 (Astronomical Society of the Pacific, Conf. Series Vol. 160, Cambridge, 1999).
19. Sulentic, J. W., Marziani, P. & Dultzin-Hacyan, D. Phenomenology of broad emission lines in active galactic nuclei. *Ann. Rev. Astron. Astrophys.* **38**, 521–571 (2000).
20. Blundell, K. M., Mioduszewski, A., Muxlow, T. W. B., Podsiadlowski, P. & Rupen, M. Images of an equatorial outflow in SS433. *Astrophys. J.* (in the press); Preprint astro-ph/0109504 at (<http://xxx.lanl.gov>) (2001).
21. Straubmeier, C., Kanbach, G. & Schrey, F. *Exp. Astron.* (in the press); Preprint astro-ph/0109191 (<http://xxx.lanl.gov>) (2001).
22. Revnivtsev, M., Sunyaev, R. & Borozdi, K. Discovery of a 0.08 Hz QPO in the black hole candidate XTE1118+480. *Astron. Astrophys.* **361**, L37–L39 (2000).

## Acknowledgements

We thank the XTE time-allocation team and Skinakas observatory for their support and flexibility in making these observations possible, and F. Schrey for technical support. We thank A. Zdziarski, K. Horne and R. Hynes for discussions. This work was done with support from the European Commission, the Training and Mobility of Researchers (TMR) research network programme.

## Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to H.S. (e-mail: [henk@mpa-garching.mpg.de](mailto:henk@mpa-garching.mpg.de)).

# Absolute-phase phenomena in photoionization with few-cycle laser pulses

G. G. Paulus\*, F. Grasbon\*, H. Walther\*, P. Villorosi†, M. Nisoli‡, S. Stagira‡, E. Priori‡ & S. De Silvestri‡

\* Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching bei München, Germany

† INFN and DEL—University of Padua, Via Gradenigo 6/a, 35131 Padua, Italy

‡ INFN, Dipartimento di Fisica, Politecnico di Milano, Piazza L. da Vinci 32, 20133 Milan, Italy

Currently, the shortest laser pulses<sup>1</sup> that can be generated in the visible spectrum consist of fewer than two optical cycles (measured at the full-width at half-maximum of the pulse's envelope). The time variation of the electric field in such a pulse depends on the phase of the carrier frequency with respect to the envelope—the absolute phase. Because intense laser–matter interactions generally depend on the electric field of the pulse, the absolute phase is important for a number of nonlinear processes<sup>2–8</sup>. But clear evidence of absolute-phase effects has yet to be detected experimentally, largely because of the difficulty of stabilizing the absolute phase in powerful laser pulses. Here we use a technique that does not require phase stabilization to demonstrate experimentally the influence of the absolute phase of a short laser pulse on the emission of photoelectrons. Atoms are ionized by a short laser pulse, and the photoelectrons are recorded with two opposing detectors in a plane perpendicular to the laser beam. We detect an anticorrelation in the shot-to-shot analysis of the electron yield.

Investigation of ultrafast processes driven by femtosecond laser pulses has acquired significance in many fields of the natural sciences. In addition, the short pulses provide access to coherent light sources with spectrally broad bandwidth and generate extremely high intensities with table-top laser systems. The key to some of these developments was the invention of laser systems delivering pulses in the range of 5 fs with energies higher than 100  $\mu$ J

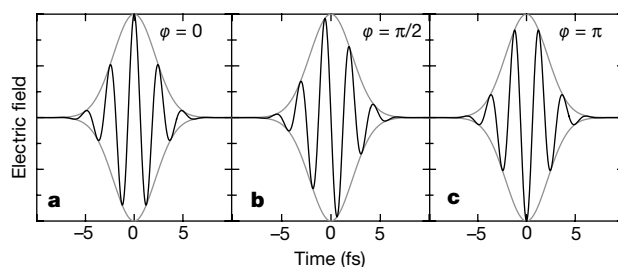
(ref. 1). The fields generated by these pulses are comparable to, or even higher than, the internal fields of atoms ( $>10^{14}$  W cm<sup>-2</sup>). Atoms are therefore easily ionized. Ultrashort pulses have the advantage that the highest pulse intensity is reached in a time shorter than that which the electron needs to escape from an atom, allowing the use of much higher effective field strengths. These facts and the need for even higher time resolution and broader bandwidth are promoting the development of shorter and shorter pulses in fields as varied as particle acceleration, plasma physics, coherent control of chemical reactions<sup>4</sup>, frequency metrology<sup>5–7</sup>, generation of coherent soft-X-ray radiation<sup>9,10</sup>, and generation of attosecond pulses<sup>8,11</sup>.

Until now, the only proven way to produce powerful laser pulses in the 5-fs regime has relied on the hollow-fibre technique<sup>1,12</sup>: pulses from a femtosecond oscillator are amplified to the millijoule level, compressed to their bandwidth limit of about 20 fs, and then fed into a hollow gas-filled fibre, where they broaden spectrally by self-phase modulation. Residual phase errors ('chirps') are then compensated when the pulses are reflected from specifically designed dielectric mirrors. The pulse repetition rate of such systems is determined by that of the pump laser of the amplifier, and is typically 1 kHz. Probably the most noteworthy property of such short pulses is that the duration of the oscillation period of the radiation (2.6 fs, corresponding to the typical Ti:sapphire femtosecond laser central wavelength of 800 nm) approaches the pulse duration. This means that such a laser pulse consists of just a few cycles, causing a new quantity to become important: the phase  $\varphi$  of the carrier with respect to the envelope of the pulse—the so-called absolute phase of the pulse. In the few-cycle regime, the electric field  $E$  should be written as

$$E(t) = E_0(t) \sin(\omega t + \varphi)$$

where  $E_0(t)$  denotes the envelope of the pulse. Examples of such pulses are shown in Fig. 1a–c. It is evident that the electric field as a function of time depends on the absolute phase, although the envelope is the same for all pulses.

Most effects in intense-field atom interaction—for example, the photoionization discussed here—depend on the time dependence of the electric field. These effects must be expected likewise to depend on the absolute phase<sup>2,3</sup>. A particularly important example is the generation of isolated attosecond pulses in the soft-X-ray spectral region<sup>8</sup>. For photoionization, a vivid picture can be obtained of the processes taking place: photoionization can be regarded as the consequence of an oscillating electric field that shakes the electrons of an atom until one or more of them gain so much energy that they can leave the atom. It is clear that the absolute phase should have an influence on this process: for very short pulses, it makes a difference whether the atom receives the first blow from one side or from the other. Unfortunately, no way has so



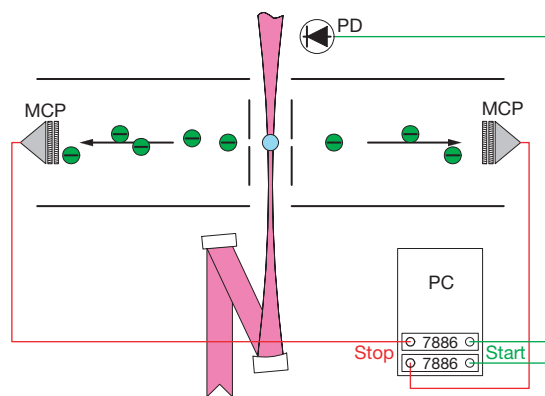
**Figure 1** The time variation of the electric field of laser pulses consisting of very few optical cycles depends on the phase  $\varphi$  of the carrier frequency with respect to the pulse's envelope. The maximum of the electric field points in opposite directions at  $t = 0$  for  $\varphi = 0$  and  $\varphi = \pi$ . The grey curves indicate the envelope of the pulses. **a**,  $\varphi = 0$ ; **b**,  $\varphi = \pi/2$ ; **c**,  $\varphi = \pi$ .

far been found to control the absolute phase in laser systems as outlined above. Rather, it must be expected to fluctuate randomly from pulse to pulse. In contrast, for femtosecond laser oscillators—that is, for pulses in the nanojoule region at repetition rates of tens of megahertz—techniques have been invented that stabilize at least the phase difference between consecutive pulses<sup>5–7,13</sup>. This is of importance to optical frequency metrology<sup>14,15</sup>.

Preliminary evidence of a regime of influence of the absolute phase on the emission of high harmonics of the laser radiation (that is, high-harmonic generation) has been reported<sup>16</sup>. Pulses of 20 fs duration were used, for which no effects of the absolute phase can usually be expected. However, high-harmonic generation is strongly governed by phase-matching effects that may increase the sensitivity under appropriate circumstances.

We present here a demonstration of the influence of the absolute phase on intense-field photoionization, investigated with a 5-fs laser system. Photoionization of atoms in intense laser fields is characterized by above-threshold ionization (ATI)<sup>17</sup>. ATI means that an atom may absorb more photons than are necessary for its ionization. This leads to photoelectron energies considerably higher than the photon energy. The effects of the absolute phase on ATI spectra (and the related phenomenon of high-harmonic generation) have been studied theoretically<sup>18–22</sup>. Typically, ATI is investigated by means of time-of-flight spectroscopy (Fig. 2).

If the interaction of atoms with long pulses is investigated, it is found that the intensity of photoelectrons emitted to opposite sides is identical, thus ensuring inversion symmetry. This is because the electric field of a long pulse (as well as the atom) exhibits inversion symmetry. This holds for any laser polarization. For short pulses, however, it can be seen from Fig. 1 that, depending on the absolute phase, the creation of photoelectrons will violate inversion symmetry. The nonlinear character of strong-field photoionization, for which the photon energy is much smaller than the ionization threshold of the atom, supports this assertion. So the absolute phase determines whether more electrons are emitted to the one or the other side. For example, a pulse as depicted in Fig. 1b can be expected to lead to an equal number of photoelectrons escaping to the left and to the right, whereas pulses like those in Fig. 1a and c will



**Figure 2** Experimental set-up. Laser pulses (6 fs full-width at half-maximum) are focused ( $f = 1$  m) onto a gas jet (indicated in blue) in a vacuum apparatus. Slits with a width of 250  $\mu\text{m}$  allow only photoelectrons originating from a predefined region in the laser focus to enter the shielded drift tubes. At the ends of the drift tubes the electrons are recorded by microchannel plates (MCP). The electrons' time of flight is measured by two computer-hosted multiscalers (FAST 7886, FAST ComTec, Oberhaching, Germany). The start signal is generated by a fast photodiode (PD), indicated on the upper part of the figure, and the stop signals are created by the MCPs. The electrons' time of flight can be used to calculate their kinetic energy. For the measurements presented in Fig. 3, the polarization of the laser pulses is changed to circular with an achromatic quarter-wave plate (B. Halle, Berlin). The waveplate is placed between the laser amplifier and the hollow-fibre compressor.

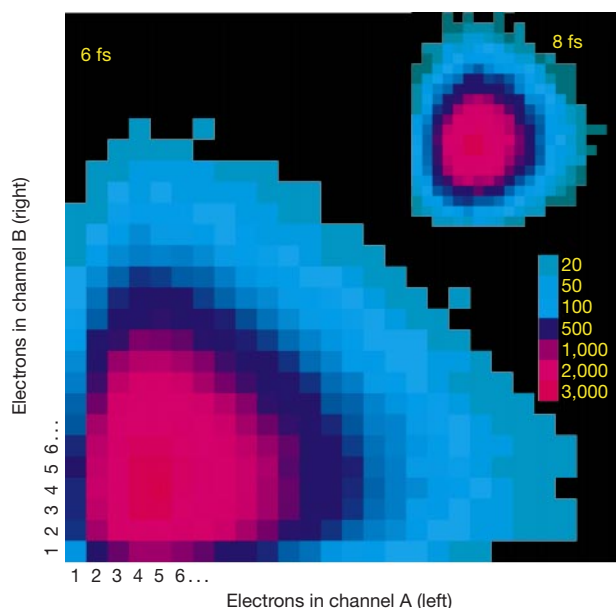
lead to asymmetric yields. Altogether, the absolute phase is expected to cause an anticorrelation in the number of electrons escaping in opposite directions.

In line with the arguments given above, the idea behind our experiment is to place two electron detectors in opposite directions with respect to the laser focus (a 'stereo ATI experiment'; Fig. 2). For each laser pulse, the number of electrons detected with both detectors is recorded. (We also measured and stored the time of flight for each electron, but this information is not used here.) Accordingly, each laser pulse is characterized by two numbers: that is, the number of electrons detected in the left-hand and that in the right-hand arm of the ATI spectrometer. We then interpret these pairs of numbers as coordinates, and accumulate thus-characterized laser shots in a map (contingency map). As the signature of an effect from the absolute phase is an anticorrelation, it should appear as a structure inclined at  $-45^\circ$  in the contingency map. This approach is expected to be very sensitive to correlations, because, in principle, an arbitrary number of laser shots can be used in order to make even small (anti)correlations statistically significant.

In implementing this idea, a number of details have to be considered in order to preserve the fragile effects of the absolute phase. This was pointed out in a theoretical analysis of the visibility of these effects in the angular distribution of the photoelectrons<sup>19</sup>. Because our correlation method is just a different—although very sensitive—way of looking at the angular distributions considered in that report, the results obtained there are applicable to our measurements. In ref. 19 it is also pointed out that the use of circularly polarized light is advantageous. The electric field of pulses with this polarization is reminiscent of a screw with a diameter that increases and decreases in accordance with the envelope of the pulse. The absolute phase determines the direction in which the maximum of the field points. For symmetry reasons, it also determines the direction in which most photoelectrons will fly. For linear polarization, the dynamics of the electrons is confined to the polarization axis, and the dependence on the absolute phase is more complicated. However, there are also disadvantages for circular polarization. The most severe one is that it is difficult to produce a circularly polarized beam, owing to the considerable bandwidth of the laser radiation. The benefits of circular polarization were therefore much smaller than expected. For circular polarization, the shortest pulse duration obtained in our experiment was 6 fs. A particularly difficult problem is presented by laser fluctuations, because they produce (positive) correlations: stronger pulses produce more electrons than weaker ones in a highly nonlinear dependence, irrespective of the emission angles of the electrons. This means that the anticorrelation produced by the absolute phase has to be stronger than the correlations due to laser-pulse fluctuations in order to be clearly measurable.

Figure 3 shows a contingency map obtained for krypton at an intensity of  $5 \times 10^{13} \text{ W cm}^{-2}$  and circular laser polarization. The signature of anticorrelation—features inclined at approximately  $-45^\circ$ —is clearly visible. In order to prove the anticorrelation and evaluate its strength and significance, we used statistical methods. One possibility is to calculate Kendall's  $\tau$  (ref. 23). A positive value of  $\tau$  indicates (positive) correlation, and a negative value, anticorrelation. Kendall's  $\tau$  is the result of a non-parametric algorithm that relies on rank-ordering: essentially, the number of concordant and discordant events is subtracted. It should be noted that non-parametric procedures, in particular Kendall's  $\tau$ , are conservative in the sense that if they indicate (anti)correlation, then it definitely exists. In our measurements with short pulses, we achieved small, negative values of  $\tau$ , (typically  $\tau \approx -0.05$ ) but of distinct significance (typically more than two standard deviations of the null hypothesis of no correlation). This was compared with the result of the same procedure for longer pulses ( $>7$  fs). There, Kendall's  $\tau$  had a positive sign with a significance of typically six standard deviations or more. This is





**Figure 3** Contingency map. Every laser shot is recorded in this plot according to the number of electrons measured in the left (channel A) and the right (channel B) arm of the electron spectrometer. In the colour code, the pixel colours represent the number of laser shots with electron numbers given by the coordinates of the pixel. The signature of the absolute phase is an anticorrelation in the number of electrons recorded with the left and the right detector. In the contingency map they appear as features inclined at  $-45^\circ$ . Shown here (main figure) is a typical measurement with krypton atoms for circular laser polarization, a pulse duration of 6 fs, and an intensity of  $5 \times 10^{13} \text{ W cm}^{-2}$ . The data were taken over 200,000 laser shots, corresponding to a measuring time of 200 s. On average, each laser shot led to the recording of approximately 5 electrons in each arm of the spectrometer. Inset, the corresponding measurement for slightly longer pulses, showing no indication of anticorrelation.

expected because of the inevitable laser pulse fluctuations, as described above. In the light of this, the anticorrelation observed is a rather strong effect.

This is, to our knowledge, the first unambiguous measurement of an effect stemming from the absolute phase of an ultrashort pulse, and we believe that the method presented here is important for its potential applications. These include detailed investigations of the influence of the absolute phase on strong-field ionization. The close relationship of ATI and coherent soft-X-ray generation (high-harmonic generation) will also make such investigations instrumental in understanding absolute-phase effects in high-harmonic generation. One way to investigate ATI in detail is to measure the strength of the anticorrelation as a function of the photoelectron's kinetic energy, the laser intensity, and so on. The strength of the anticorrelation can also serve as a measure of the pulse duration, which is useful because such measurements become increasingly difficult for shorter and shorter pulses when conventional methods are used. Finally, with suitable refinements, our stereo ATI apparatus could be used to measure the absolute phase. The measured value could then be used to interpret other experiments, such as the generation of coherent soft X-rays, generation of isolated attosecond laser pulses, or coherent control of chemical reactions. □

Received 27 July; accepted 17 September 2001.

1. Nisoli, M. *et al.* Compression of high-energy laser pulses below 5 fs. *Opt. Lett.* **22**, 522–524 (1997).
2. Brabec, T. & Krausz, F. Intense few-cycle laser fields: frontiers of nonlinear optics. *Rev. Mod. Phys.* **72**, 545–591 (2000).
3. Krausz, F. From femtochemistry to attophysics. *Phys. World* **14**, 41–46 (2001).
4. Rabitz, H., de Vivie-Riedle, R., Motzkus, M. & Kompa, K. Whither the future of controlling quantum phenomena? *Science* **288**, 824–828 (2000).

5. Reichert, J. *et al.* Measuring the frequency of light with mode-locked lasers. *Opt. Commun.* **172**, 59–68 (1999).
6. Jones, D. J. *et al.* Carrier-envelope phase control of femtosecond mode-locked lasers and direct optical frequency synthesis. *Science* **288**, 635–639 (2000).
7. Apolonski, A. *et al.* Controlling the phase evolution of few-cycle light pulses. *Phys. Rev. Lett.* **85**, 740–743 (2000).
8. Drescher, M. *et al.* X-ray pulses approaching the attosecond frontier. *Science* **291**, 1923–1927 (2001).
9. Spielmann, Ch. *et al.* Generation of coherent X-rays in the water window using 5-femtosecond laser pulses. *Science* **278**, 661–664 (1997).
10. Rundquist, A. *et al.* Phase-matched generation of coherent soft X-rays. *Science* **280**, 1412–1415 (1998).
11. Paul, P. M. *et al.* Observation of a train of attosecond pulses from high harmonic generation. *Science* **292**, 1689–1692 (2001).
12. Nisoli, M., De Silvestri, S. & Svelto, O. Generation of high energy 10 fs pulses by a new pulse compression technique. *Appl. Phys. Lett.* **68**, 2793–2795 (1996).
13. Telle, H. R. *et al.* Carrier-envelope offset phase control: A novel concept for absolute optical frequency and ultrashort pulse generation. *Appl. Phys. B* **69**, 327–332 (1999).
14. Eckstein, J. N. *High-resolution Spectroscopy using Multiple Coherent Interactions*. Thesis, Stanford Univ. (1978).
15. Wineland, D. J., Bergquist, J. C., Itano, W. M., Diedrich, F. & Weimer, C. S. in *The Hydrogen Atom* (eds Bassani, G. F., Inguscio, M. & Hänsch, T. W.) 123–133 (Springer, Berlin, 1989).
16. Durfee, Ch. G. *et al.* Phase matching of high-order harmonics in hollow waveguides. *Phys. Rev. Lett.* **83**, 2187–2190 (1999).
17. Agostini, P., Fabre, F., Mainfray, G., Petite, G. & Rahman, N. Free-free transitions following six-photon ionization of xenon atoms. *Phys. Rev. Lett.* **42**, 1127–1130 (1979).
18. Cormier, E. & Lambropoulos, P. Effect of the initial phase of the field in ionization by ultrashort laser pulses. *Euro. Phys. J. D* **2**, 15–20 (1998).
19. Dietrich, P., Krausz, F. & Corkum, P. B. Determining the absolute carrier phase of a few-cycle laser pulse. *Opt. Lett.* **25**, 16–18 (2000).
20. Christov, I. P. Phase-dependent ionization in the barrier suppression regime. *J. Appl. Phys.* **70**, 459–462 (2000).
21. de Bohan, A., Antoine, Ph., Milosevic, D. B. & Piraux, B. Phase-dependent harmonic emission with ultrashort laser pulses. *Phys. Rev. Lett.* **42**, 1127–1130 (1979).
22. Potvliege, R. M., Klystra, N. J. & Joachain, C. J. Photon emission by  $\text{He}^+$  in intense ultrashort laser pulses. *J. Phys. B* **33**, L743–L748 (2000).
23. Press, W. P., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. *Numerical Recipes* (Cambridge Univ. Press, Cambridge, 1986).

## Acknowledgements

We thank A. Kasper, H. Stehbeck and K. Witte for their collaboration in the early stage of this project.

Correspondence and requests for materials should be addressed to G.G.P. (e-mail: ggp@mpq.mpg.de).

## Sharper images by focusing soft X-rays with photon sieves

L. Kipp\*, M. Skibowski\*, R. L. Johnson†, R. Berndt\*, R. Adelung\*, S. Harm\* & R. Seemann‡

\* Institut für Experimentelle und Angewandte Physik, Universität Kiel, D-24098 Kiel, Germany

† Institut für Experimentalphysik, Universität Hamburg, D-22761 Hamburg, Germany

‡ Niedmers & Seemann, European Patent Attorneys, D-22767 Hamburg, Germany

Fresnel zone plates consisting of alternating transmissive and opaque circular rings can be used to focus X-rays<sup>1</sup>. The spatial resolution that can be achieved with these devices is of the order of the width of the outermost zone and is therefore limited by the smallest structure (20–40 nm) that can be fabricated by lithography today<sup>2</sup>. Here we show that a large number of pinholes distributed appropriately over the Fresnel zones make it possible to focus soft X-rays to spot sizes smaller than the diameter of the smallest pinhole. In addition, higher orders of diffraction and secondary maxima can be suppressed by several orders of magnitude. In combination with the next generation of synchrotron

light sources (free-electron lasers) these 'photon sieves' offer new opportunities for high-resolution X-ray microscopy and spectroscopy in physical and life sciences.

A variety of optical devices have been developed to focus X-ray radiation<sup>3</sup>, ranging from X-ray mirrors to concentrators in the form of tapered capillaries<sup>4</sup>, compound refractive lenses<sup>5</sup> and Fresnel zone plates<sup>6</sup>. The most severe limitation in using refracting optics for focusing electromagnetic radiation arises from absorption. The visible spectral region has a large number of optical materials available for fabricating lenses. At shorter wavelengths all solids are strongly absorbing until the hard X-ray region (10–50 keV). Recently, compound refractive lenses for use with hard X-rays have been fabricated—consisting of tens or hundreds of small holes, or hollow spheres, arranged in arrays to achieve focusing in one or two dimensions<sup>5</sup>. In the vacuum ultraviolet (VUV) and extreme ultraviolet (XUV) regime (10–1 keV) strong absorption rules out this technique but Fresnel zone plates provide an effective solution<sup>6</sup>. Nevertheless, improvements are desirable in three areas: (1) overcoming the limitation in the spatial resolution given by the width of the zone-plate outermost zone; (2) elimination of unwanted diffraction orders; (3) reduction of the scattered intensity arising from the edge of the finite-sized zone plate.

Here we demonstrate that an appropriate distribution of pinholes over the Fresnel zones (a photon sieve) can be used to focus soft X-rays effectively. In contrast to conventional zone-plate optics, a sharpening of the focal spot combined with an effective suppression of higher orders and finite-size effects is achieved by exploiting variations in pinhole diameter and pinhole distribution to optimize the optical performance.

In the standard optical configuration shown in Fig. 1, a photon sieve with a quasi-random distribution of transmissive pinholes (white) replaces a comparable conventional zone plate which consists of transmissive (grey) and opaque (black) rings. Designing a photon sieve with a given smallest pinhole diameter  $d_{\min}$  is simple. We consider point-to-point focusing of a source at finite distance  $p$  from a photon sieve to a focus at distance  $q$ . For simplicity, the source and image are on the optical axis (in general, however, photon sieves can be constructed where the source or image point, or both, are off-axis). To obtain a distinct first-order focus, the pinholes have to be positioned such that the optical path length from the source via the centre of the pinholes to the focal point is an integral number of wavelengths  $\lambda$ . Consequently, the pinholes have

to be centred at distances  $r_n$  from the optical axis given by:

$$\sqrt{r_n^2 + p^2} + \sqrt{r_n^2 + q^2} = p + q + n\lambda \quad (1)$$

where  $\lambda$  is the wavelength of the radiation and  $n$  a positive integer. By choosing random numbers for  $n$  and  $\varphi$  ( $0 < \varphi \leq 2\pi$ ), an uncorrelated distribution of pinholes with centres at  $(r_n, \varphi)$  can be generated. A useful feature of our diffraction optics is that aberrations can be corrected by modifying both the positions of the pinholes and their distribution.

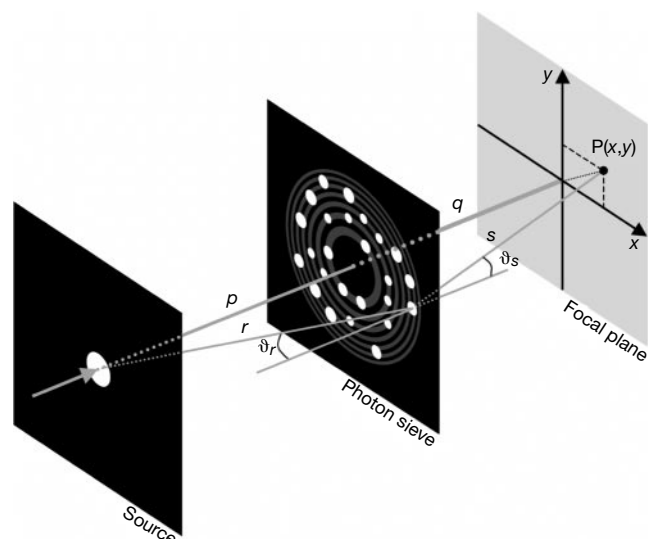
Next we discuss how to choose the diameters of the pinholes. The trivial case is for pinholes of diameter  $d < w$  centred on a white (constructive interference) zone of width  $w$  (Fig. 2a). Light from all parts of the pinhole interferes constructively at the focus. Pinholes with diameters larger than the width of the underlying zone can also be used (Fig. 2b). Such holes transmit contributions that partially compensate—indicated by the black and white areas. As long as the white areas dominate there is a net positive contribution to the focus. If the black contributions were to predominate (Fig. 2c) the pinhole just needs to be centred on a black zone (Fig. 2d) instead of a white zone (Fig. 2c) to produce again a positive contribution to the focus.

We calculate the light amplitude at  $P$  (see Fig. 1) using Fresnel–Kirchhoff diffraction theory

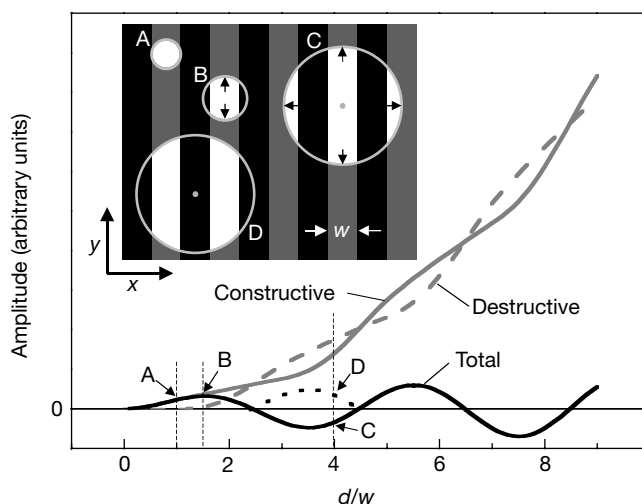
$$E(P) = -i \frac{A}{2\lambda} \int_{\text{holes}} \frac{e^{ik(r+s)}}{rs} (\cos \vartheta_r + \cos \vartheta_s) dS \quad (2)$$

where  $A$  is the amplitude at unit distance from the source and the integral is performed over the pinholes.

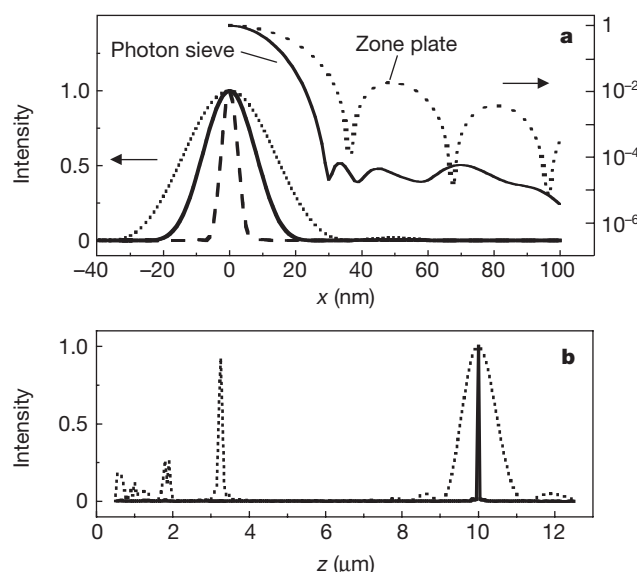
Calculated contributions to the focal amplitude of a single pinhole as a function of  $d/w$  are plotted in Fig. 2. Contributions from the white (constructive) and black (destructive) parts are calculated separately. Considering first only the white regions, we observe for small pinhole diameters  $d < w$  a quadratic increase of constructive contributions to the focal amplitude until the pinhole diameter equals the width  $w$  of the underlying transmissive Fresnel zone ( $A$  in Fig. 2). For larger  $d$  the white area in the  $y$  direction increases essentially linearly (see arrows in B in Fig. 2 inset), but for  $d > 3w$  the increase is nominally quadratic because the white area also



**Figure 1** Diagram showing point-to-point imaging with a diffractive optical element. The details are described in the text.



**Figure 2** Examples of photon sieve pinholes together with the underlying zone-plate geometry and the contributions of a single pinhole to the focal amplitude. Constructive and destructive interference contributions from 'white' and 'black' areas inside the pinhole are plotted as solid and dashed grey curves, respectively. The total amplitude at the focus from a single pinhole of diameter  $d$  centred on a transmissive zone (A, B, C) of width  $w$  is given by the solid black curve; the total contribution to the amplitude at the focus from a pinhole centred on an opaque zone (D) is given by the dotted black curve.



**Figure 3** Comparison of calculated intensity distributions of photon sieves and zone plates with smallest structure sizes  $d_{\min} = w_{\min} = 30$  nm. **a**, Transverse intensity distribution in focal plane (left axis shows a linear scale, right axis shows a log scale) for a zone plate (dotted line) with a rectangular transmission window and two different photon sieves:  $(d/w)_{\max} = 2.4$  (solid line),  $(d/w)_{\max} = 7.5$  (dashed line),  $\lambda = 2.4$  nm,  $p = 20$  m,  $q = 500$   $\mu$ m. For the photon sieves (95,567 pinholes) the pinhole density on each ring is

modulated to produce a Weber-type transmission window. For clarity the dashed curve is not shown in the log plot. **b**, Longitudinal intensity distribution on the optical axis of a zone plate (dotted line) and a photon sieve (solid line, 13,440 pinholes):  $\lambda = 6.2$  nm,  $p = \infty$ ,  $q = 10$   $\mu$ m. Calculations were performed for a bandwidth  $\Delta\lambda/\lambda = 10^{-4}$ . Intensities are normalized to the peak intensity.

increases in the  $x$  direction (C in Fig. 2). The constructive contribution to the focal amplitude is indicated by the solid grey curve.

Competing with the increasing constructive contribution (white areas) is an increase of destructive contributions (black areas within the pinhole) as indicated by the dashed grey curve in Fig. 2. For  $d < w$  this contribution is zero; it then increases quadratically up to  $d = 3w$ , linearly to  $d = 5w$ , and so on.

The resulting oscillating total contribution to the focal amplitude is given by the solid black curve. Constructive and destructive contributions compensate at  $d/w$  values of about 2.4, 4.4, 6.4, and so on. The largest focal amplitudes from a single pinhole are for pinhole diameters  $d$  of about  $1.5w$ ,  $3.5w$ ,  $5.5w$ , and so on (corresponding to 1st, 3rd, 5th, ... order of diffraction at the focus). The sign of the focal amplitude is reversed when pinholes are centred on opaque zones instead of transmissive zones, so pinholes with diameters around  $3.5w$ ,  $7.5w$ , and so on, should be centred on an opaque zone to give positive contributions to the focal amplitude (D in Fig. 2). Combining pinholes corresponding to different maxima allows mixed-order photon sieves to be constructed in analogy with composite zone plates<sup>7</sup> which also use combined orders of diffraction.

The ultimate spatial resolution  $\Delta x$  of a conventional zone plate is limited by the width  $w_{\min}$  of the outermost zone:  $\Delta x \approx w_{\min}$  (first order of diffraction); when working in higher orders of diffraction  $m$  (odd) smaller spots ( $\sim 1/m$ ) with reduced intensity ( $\sim 1/m^2$ ) can also be achieved. For a photon sieve, however, even in the first order of diffraction the spatial resolution can be smaller than the smallest pinhole diameter  $d_{\min}$ . In fact, it is limited by the width of the outermost zone of the underlying zone plate geometry. This effective smallest width  $w_{\min}^{\text{eff}}$  depends on the maximum  $d/w$  ratio employed:

$$\Delta x \approx w_{\min}^{\text{eff}} = d_{\min} \left/ \left( \frac{d}{w} \right)_{\max} \right. \quad (3)$$

Calculated spot sizes for a zone plate (dotted curves) and photon sieves (solid and dashed curves) are shown in Fig. 3a. For all three diffractive optical elements the smallest structure size is 30 nm.

Focal spot sizes (full-width at half-maximum, FWHM) are 32 nm for the zone plate (in first order) and 18 nm and 6 nm for the photon sieves with  $(d/w)_{\max} = 2.4$  (in first order) and  $(d/w)_{\max} = 7.5$  (in 1st + 3rd + 5th + 7th order), respectively. The highest spatial resolution achievable with photon sieves is limited by the accuracy in positioning the centre of a pinhole on the corresponding zone. With today's fabrication procedures an accuracy of 2 nm can be achieved<sup>8</sup>, which is sufficient to place 30 nm pinholes on a 6 nm effective zone.

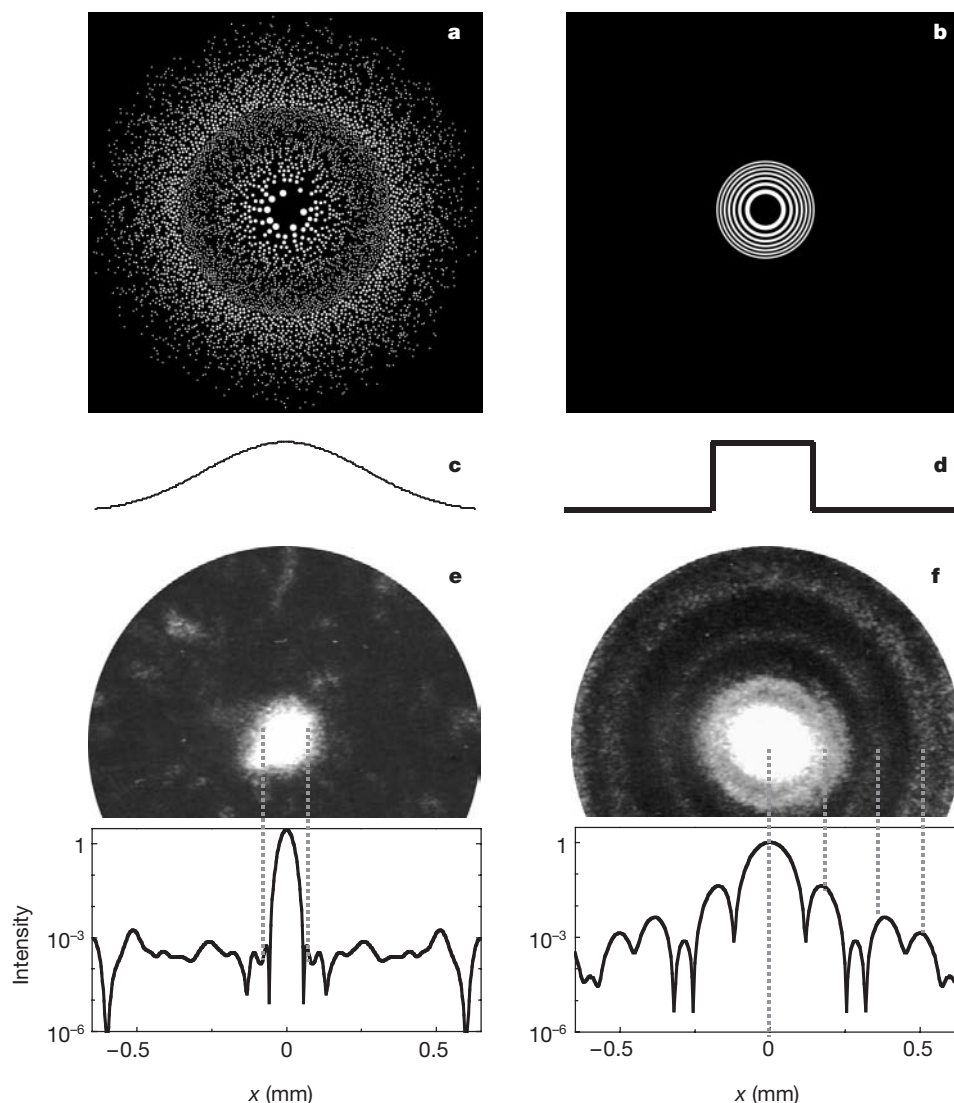
It should be noted that to achieve a potentially better spatial resolution of mixed-order photon sieves or composite zone plates a bandwidth of the light source  $\Delta\lambda/\lambda < 10^{-4}$  is required. Photon sieves with pinhole diameters  $d < 2.4w$ , however, which already increase the spatial resolution to  $\Delta x = d_{\min}/2.4$  exclusively employ first-order diffraction and do not need a high degree of monochromaticity.

In addition to the principal focus, zone plates also produce ring-shaped secondary maxima that decrease the signal-to-noise ratio and blur the images. This phenomenon is analogous to the production of side bands in a digital filter with a rectangular transmission window. Optimized window functions are routinely used in digital filtering<sup>9</sup>. By adapting this concept, we have implemented a Weber-type transmission window by modulating the pinhole density on each ring of the photon sieve. In this way, as shown in Fig. 3a, we can achieve a substantial suppression of the secondary maxima. A similar concept (for example, adjusting the mark-to-space ratio) could be implemented for zone plates.

Higher-order diffraction occurs when the diffractive optical element has the appropriate symmetry. The higher (odd) orders along the optical axis of a Fresnel zone plate are well known (Fig. 3b, dotted line). With a photon sieve (Fig. 3b, solid line) such higher-order contributions can be efficiently reduced by the random distribution of the pinholes on each zone.

A direct comparison of focal intensities of photon sieves and zone plates is difficult because of the possibility of using higher-order diffraction and transmission windows, for example. To give a basic idea, we compare first-order diffractive optics of equal diameter. A zone plate has a transmission of 50% whereas a photon sieve





**Figure 4** Optical prototypes for a photon sieve and a zone plate together with experimental and calculated intensity distributions. **a**, Photon sieve (5,646 pinholes,  $(d/w)_{\max} = 4$ ) with Weber-type transmission window **(c)**. **b**, Zone plate with rectangular transmission window **(d)**. The smallest structure sizes in both cases are  $100\ \mu\text{m}$

( $\lambda = 632.8\ \text{nm}$  (He–Ne laser),  $p = 20\ \text{m}$ ,  $q = 1\ \text{m}$ ). **e**, **f**, Photographs of focal plane intensities compared to calculated intensities. Intensities are normalized to the zone-plate peak intensity.

transmits only 15–20% of the incident light. The intensity scales as the square of the area, so the first-order diffraction intensity is lower than that of a zone plate of equal diameter by a factor of 10. However, despite this lower intensity, experimentation with photon sieves is possible on acceptable timescales with present-day synchrotron light sources.

For an experimental verification of the above results we have constructed a photon sieve and a zone plate with a minimum structure size of  $100\ \mu\text{m}$  working in the optical region at  $632.8\ \text{nm}$  wavelength (He–Ne laser). The diffraction optics depicted in Fig. 4a and b were printed with 4,000 lines on 35-mm slides. The transmission windows used are plotted in Fig. 4c and d. Photographs of the intensity patterns on the focal plane are compared to calculated focal-plane intensities in Fig. 4e and f. We note good agreement between experimental and theoretical results concerning peak widths and reduction of secondary maxima.

Photon sieves can be produced using techniques similar to those for making zone plates. The limitations of conventional zone plates can be overcome by using pinholes with varying diameters instead of concentric rings as the basic diffracting elements. This

generalization enables the development of a new class of diffractive optical elements with additional design parameters that can result in unprecedented focusing of electromagnetic radiation to spot sizes well below the dimensions of the smallest diffracting structure. The development of fourth-generation synchrotron light sources (free-electron lasers) that deliver transversally coherent radiation in the soft X-ray regime in combination with photon sieves will provide a wealth of new opportunities in X-ray microscopy, spectroscopy and lithography. Photon sieves for use at the Hamburg Free-Electron Laser facility (HASYLAB)/DESY synchrotrons are being developed.

Received 30 May; accepted 4 September 2001.

1. Thieme, J., Schmahl, G., Rudolph, D. & Umbach, E. *X-Ray Microscopy and Spectroscopy* IV-3–IV-110 (Springer, Berlin, 1996).
2. Anderson, E. H., Boegli, V. & Muray, L. P. Electron beam lithography digital pattern generator and electronics for generalized curvilinear structures. *J. Vac. Sci. Technol. B* **13**, 2529–2534 (1995).
3. Attwood, D. *Soft X-Rays and Extreme Ultraviolet Radiation* (Cambridge Univ. Press, Cambridge, 2000).
4. Bilderback, D. H. & Thiel, D. J. Microbeam generation with capillary optics. *Rev. Sci. Instrum.* **66**, 2059–2063 (1995).

5. Snigirev, A., Kohn, V., Snigireva, I. & Lengeler, B. A compound refractive lens for focusing high-energy X-rays. *Nature* **384**, 49–51 (1996).
6. Schmahl, G., Rudolph, D., Guttman, P. & Christ, O. in *Zone Plates for X-ray Microscopy* Vol. 43 *X-ray Microscopy* (eds Schmahl, G. & Rudolph, D.) 63–74 (Springer Series in Optical Sciences, Academic, London, 1984).
7. Michette, A. G. *Optical Systems for Soft X-Rays* 193–194 (Plenum, New York, 1986).
8. Spector, S. J., Jacobsen, C. J. & Tennant, D. M. Process optimization for production of sub-20 nm soft X-ray zone plates. *J. Vac. Sci. Technol. B* **15**, 2872–2876 (1997).
9. Cappellini, V., Constantinides, A. G. & Emiliani, P. *Digital Filters and their Applications* Vol. 4 *Techniques of Physics* (eds March, N. H. & Daglish, H. N.) 65–74 (Springer, Berlin, 1981).

# Acknowledgements

Discussions with B. Niemann and G. Schmahl are gratefully acknowledged. This work was supported in part by the Bundesministerium für Bildung und Forschung (BMBF).

Correspondence and requests for material should be addressed to L.K.  
(e-mail: KIPP@physik.uni-kiel.de).

## Water conduction through the hydrophobic channel of a carbon nanotube

G. Hummer\*, J. C. Rasaiah\*† & J. P. Noworyta†

\* Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, USA

† Department of Chemistry, University of Maine, Orono, Maine 04469, USA

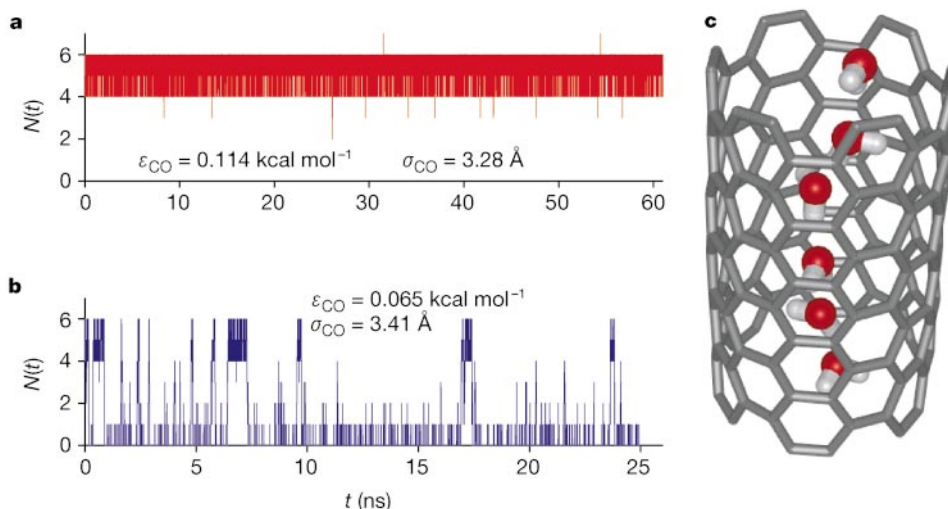
Confinement of matter on the nanometre scale can induce phase transitions not seen in bulk systems<sup>1</sup>. In the case of water, so-called drying transitions occur on this scale<sup>2–5</sup> as a result of strong hydrogen-bonding between water molecules, which can cause the liquid to recede from nonpolar surfaces to form a vapour layer separating the bulk phase from the surface<sup>6</sup>. Here we report molecular dynamics simulations showing spontaneous and continuous filling of a nonpolar carbon nanotube with a one-dimensionally ordered chain of water molecules. Although the molecules forming the chain are in chemical and thermal equilibrium with the surrounding bath, we observe pulse-like transmis-

sion of water through the nanotube. These transmission bursts result from the tight hydrogen-bonding network inside the tube, which ensures that density fluctuations in the surrounding bath lead to concerted and rapid motion along the tube axis<sup>7–9</sup>. We also find that a minute reduction in the attraction between the tube wall and water dramatically affects pore hydration, leading to sharp, two-state transitions between empty and filled states on a nanosecond timescale. These observations suggest that carbon nanotubes, with their rigid nonpolar structures<sup>10,11</sup>, might be exploited as unique molecular channels for water and protons, with the channel occupancy and conductivity tunable by changes in the local channel polarity and solvent conditions.

We designed a short<sup>12</sup> uncapped, single-walled nanotube 13.4 Å long with a diameter of 8.1 Å, and simulated for 66 ns the dynamics of this nanotube solvated in a water reservoir. Despite its strongly hydrophobic character, the initially empty central channel of the nanotube is rapidly filled by water from the surrounding reservoir, and remains occupied by about five water molecules during the entire 66 ns (Fig. 1a). These water molecules form a hydrogen-bonded chain (Fig. 1c). The number of water molecules,  $N$ , fluctuates between 2 and 7, with  $N = 2$  occurring only once in 66 ns (Figs 1a and 2a). The radial and axial density profiles of the water in the nanotube show considerable structure and density depletion at the nanotube openings (Figs 2b and c).

Water molecules entering the nanotube lose on average two out of four hydrogen bonds. Only a fraction of the lost energy ( $\sim 10$  kcal mol<sup>-1</sup>) can be recovered through van der Waals interactions with the carbon atoms of the nanotube ( $\sim 4$  kcal mol<sup>-1</sup>), while electrostatic interactions with water molecules beyond the nanotube wall are found to be negligible. Considering this loss of hydrogen bonding, and the weak attraction of water to the nanotube carbon atoms, with a Lennard–Jones well depth of only about 0.114 kcal mol<sup>-1</sup>, this persistent hydration of the nanotube interior seems surprising, but is consistent with the experimentally inferred adsorption of water onto nanotubes<sup>13</sup>. Further support comes from recent simulations<sup>14</sup> showing that a potassium iodide melt would be sucked into carbon nanotubes to form nano-crystallites, despite favourable solvation in the surrounding liquid.

The water occupancy of the channel is determined by the local excess chemical potential,  $\mu_{\text{nt}}^{\text{ex}}$ , defined as the negative free energy of removing a water molecule from the channel. This free energy is dominated not by how strongly bound a water molecule is on average, but by how populated weakly bound states are (see Methods, equation (2)). The average binding

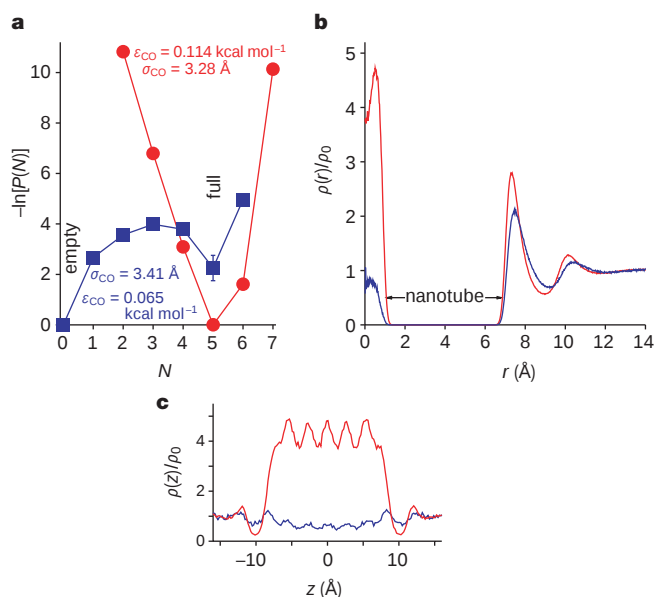


**Figure 1** Water occupancy. **a**, **b**, Number  $N$  of water molecules inside the nanotube as a function of time for  $sp^2$  carbon parameters (**a**) and reduced carbon–water attractions (**b**). **c**, Structure of the hydrogen-bonded water chain inside the nanotube.

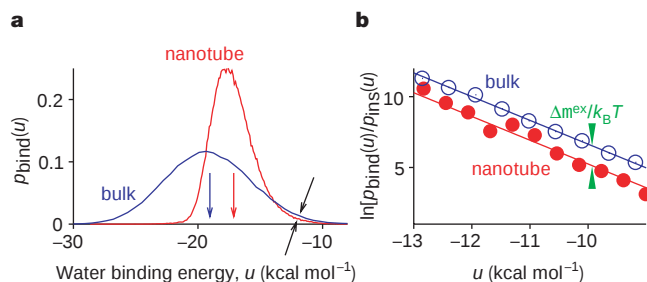
energy of water molecules inside the nanotube is indeed unfavourable compared to bulk water. Nevertheless, the binding energies inside the nanotube are more sharply distributed (Fig. 3a), and high-energy states dominating the free energy are less frequently occupied. As a consequence, although water molecules at the centre of the nanotube (within 0.8 Å of the cylinder axis) lose on average 2 kcal mol<sup>-1</sup> in energy, they have a lower excess chemical potential of  $\mu_{\text{nt}}^{\text{ex}} \approx -6.87 \pm 0.07$  kcal mol<sup>-1</sup>, compared to bulk TIP3P water ( $\mu_{\text{w}}^{\text{ex}} \approx -6.05 \pm 0.02$  kcal mol<sup>-1</sup>). The resulting chemical-potential difference (Fig. 3b) of  $\mu_{\text{nt}}^{\text{ex}} - \mu_{\text{w}}^{\text{ex}} \approx -0.82 \pm 0.09$  kcal mol<sup>-1</sup> agrees with the estimate from the average occupancy number,  $-k_{\text{B}}T \ln(\langle N \rangle / \rho_0 \Delta V) = -0.87$  kcal mol<sup>-1</sup> (see Methods, equation (1)).

Hydrogen bonds inside the nanotube are shielded from fluctuations in the environment. Only 0.02% of water pairs in contact (within 3.5 Å oxygen-atom distance) are unbound with a pair energy  $u_{ij} > 0$ , compared to 15% in bulk water. Hydrogen bonds in the nanotube are highly oriented, with less than 15% of the H-O...O angles exceeding 30°, compared to 37% in bulk water. The average lifetime of a hydrogen bond<sup>15</sup> (oxygen distance  $\leq 3.5$  Å, hydrogen-bond angle  $\leq 30^\circ$ ) is 5.6 ps, compared to 1.0 ps for bulk water. OH bonds involved in hydrogen bonds are nearly aligned with the nanotube axis, collectively flipping direction<sup>16</sup> every 2–3 ns on average. Despite this quasi-one-dimensional order, the water chain retains considerable entropy: water molecules can rotate freely about their aligned hydrogen bonds, resulting in a degenerate energetic ‘ground state’ and the narrow distribution of binding energies shown in Fig. 3a.

Water molecules not only penetrate into, but are also conducted through, the nanotube. During the 66 ns, 1,119 water molecules entered the nanotube on one side and left on the other side, corresponding to an average of about 17 water molecules per



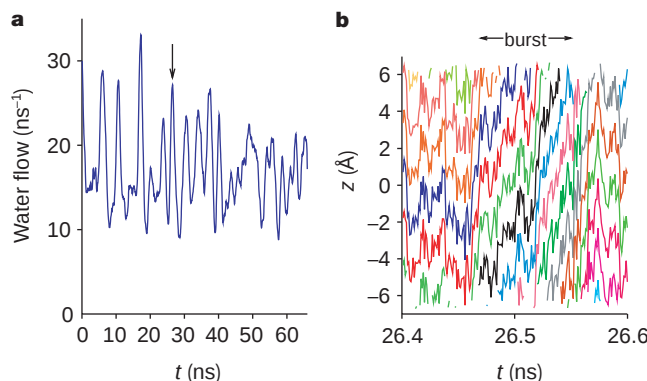
**Figure 2** Energetics and structure of nanotube water. **a**, Free energy of occupancy fluctuations,  $\beta F(N) = -\ln p(N)$ , where  $p(N)$  is the probability of finding exactly  $N$  water molecules inside the nanotube. The approximately gaussian occupancy fluctuations<sup>28</sup> of the filled nanotube with  $sp^2$  carbon parameters (red circles) become bi-modal in a predominantly empty nanotube with reduced carbon–water attractions (blue squares; see Methods). **b**, Radial water density profile in units of the bulk density  $\rho_0$ , averaged over cylindrical shells centred at the axis of the nanotube, with radius  $r$ , and height determined by the carbon atoms at the nanotube rim (red,  $sp^2$  carbon parameters; blue, modified carbon–water interactions). **c**, Water density along the nanotube axis, within a distance of 0.8 Å from the axis.



**Figure 3** Water binding energies. **a**, Probability distribution  $p_{\text{bind}}(u)$  of binding energies  $u$  for bulk water (blue) and water inside the nanotube (red; with a cylindrical volume of radius 0.8 Å from the nanotube axis and a height of about 13.4 Å). Vertical arrows indicate average binding energies. Tilted black arrows indicate the cross-over region, in which weakly bound states are more populated in bulk water. **b**,  $\ln[p_{\text{bind}}(u)/p_{\text{bulk}}(u)]$  for water inside the nanotube (filled circles) and in bulk TIP3P water (open circles), fitted to  $\beta(\mu_{\text{w}}^{\text{ex}} - u)$  (lines). The vertical distance between the parallel lines of slope  $-\beta$  gives the difference in the excess chemical potentials,  $\beta(\mu_{\text{w}}^{\text{ex}} - \mu_{\text{nt}}^{\text{ex}})$ , as indicated by green arrows.

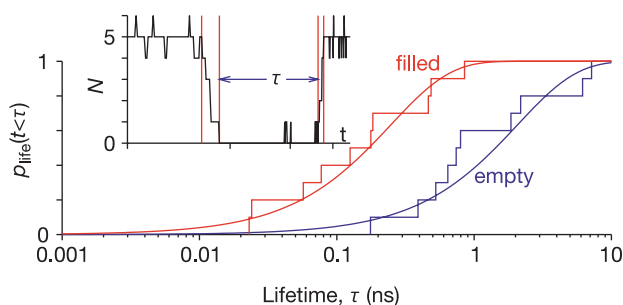
nanosecond passing through the nanotube ( $51 \times 10^{-14}$  cm<sup>3</sup> s<sup>-1</sup>). The measured water flow through the twice as long channel of the transmembrane protein aquaporin-1 is of comparable magnitude<sup>17</sup>. Water conduction occurs in pulses through the nanotube with peaks of about 30 water molecules per nanosecond (Fig. 4a), reminiscent of single ion channel activity<sup>18</sup>. Slow fluctuations of the occupancy number have been observed by simulation in the pore region of a synthetic transmembrane channel<sup>19</sup>. The burst-like water conduction follows from the tight hydrogen-bond network inside the nanotube. Rupturing the water chain is energetically costly, and thus rare. Fluctuations outside the nanotube lead to highly concerted, yet rapid, drift-like motion of the water molecules along the nanotube axis<sup>7–9</sup>, resulting in bursts in the water flow (Fig. 4). During those bursts, the water chain moves with little resistance through the ‘greasy’ nanotube, unhindered by interactions with the hydrophobic wall (Fig. 4b).

To explore the role of attractive interactions<sup>20,21</sup>, we reduce the depth of the carbon–water van der Waals potential well by 0.05 kcal mol<sup>-1</sup> (see Methods). Modifying the attractive potential mimics changes in the local polarity and solvent conditions. This seemingly small perturbation of the nanotube drastically affects its water occupancy, which fluctuates in sharp transitions between empty and filled states during a 25-ns simulation (Fig. 1b). Intermediate states are rarely populated (Fig. 2a), reflecting the large



**Figure 4** Flow of water through the nanotube. **a**, Number of water molecules leaving the nanotube at time  $t$  that entered the nanotube from the other side. Individual water conduction events are smoothed with a 1-ns-wide (at half-maximum) triangular filter. **b**, Motions  $z(t)$  of individual water molecules shown as coloured lines, inside the nanotube along the nanotube axis during the conduction ‘burst’ at 26.5 ns.





**Figure 5** Kinetics of emptying and filling transitions. Cumulative distributions  $P_{\text{life}}(t < \tau)$  of the lifetimes  $\tau$  of the filled (red) and empty (blue) states, defined as the times between the first and last state with  $N = 5$  ( $N = 0$ ) before a transition (inset). Smooth curves are exponential distributions with the same mean.

energetic cost of fragmenting the hydrogen-bonded water chain, and leading to two-state kinetics despite the small volume. The lifetimes of the filled and empty states follow exponential distributions, with means of  $\tau_{\text{filled}} = 247$  ps and  $\tau_{\text{empty}} = 2.065$  ps (Fig. 5). Kinetic and equilibrium free-energy differences between filled and empty states agree,  $\beta \Delta F_{\text{kin}} = \ln(\tau_{\text{empty}}/\tau_{\text{filled}}) \approx 2.1 \pm 0.5$ , and  $\beta \Delta F_{\text{equil}} = -\ln[\sum_{N=0}^3 p(N)/\sum_{N>3} p(N)] \approx 2.1 \pm 0.5$  (Fig. 2a).

We conclude that, counter to intuition, hydrophobic channels can have significant water occupancy despite a reduction in the number of hydrogen bonds compared to the bulk fluid. Small changes in the nanotube–water interactions can lead to large changes in the water occupancy of the channel, with two-state transitions between empty and filled states. This might have biological significance, and offer an explanation for crystallographically empty, yet functionally filled, hydrophobic channels in proton-transferring proteins<sup>22</sup>. Changes in amino-acid ionization states could trigger water influx and establish protonic connectivities. Emptying/filling transitions in membrane-inserted, functionalized nanotubes, driven, for example, by light-induced excitation of covalently bound dye molecules, could potentially be used in light sensors as single-molecule ‘field-effect transistors’ for protonic currents. □

## Methods

### Molecular dynamics simulations

The 144-carbon (6,6) nanotube (13.4 Å long and 8.1 Å in diameter) is formed by folding a graphite sheet of  $5 \times 12$  carbon rings to a cylinder. Solvated in a cubic box with 1,034 TIP3P water molecules<sup>23</sup>, the nanotube is free to translate and rotate. Molecular dynamics simulations were performed at constant pressure<sup>24</sup> (1 bar, box length  $L = 32.06 \pm 0.04$  Å) and temperature (300 K) with AMBER 6.0 (University of California at San Francisco), with particle-mesh Ewald electrostatics<sup>25</sup> and cubic-spline interpolation ( $\sim 1$  Å grid width). A time step of 2 fs was used. Structures were saved every 1 ps. The carbon atoms are modelled as uncharged Lennard–Jones particles with a cross-section of  $\sigma_{\text{CC}} = 3.400$  Å and a depth of the potential well of  $\epsilon_{\text{CC}} = 0.086$  kcal mol<sup>−1</sup>, corresponding to  $sp^2$  carbons in the AMBER96 force field<sup>26</sup>. Carbon–carbon bond lengths of 1.4 Å and bond angles of 120° are maintained by harmonic potentials with spring constants of 938 kcal mol<sup>−1</sup> Å<sup>−2</sup> and 126 kcal mol<sup>−1</sup> rad<sup>−2</sup>. In addition, a weak dihedral angle potential is applied to bonded carbon atoms. The carbon–water Lennard–Jones parameters are  $\sigma_{\text{CO}} = 3.2751$  Å and  $\epsilon_{\text{CO}} = 0.114333$  kcal mol<sup>−1</sup>. We also conducted a 25-ns simulation with modified carbon–water Lennard–Jones interactions of  $\sigma_{\text{CO}} = 3.4138$  Å and  $\epsilon_{\text{CO}} = 0.06461$  kcal mol<sup>−1</sup>, thus weakening the water–nanotube van der Waals attractions by a factor of about two. For reference, bulk TIP3P water was simulated for 3 ns (256 molecules, 300 K at a density of 0.987 g cm<sup>−3</sup>).

### Water occupancy and chemical potentials

The average number  $\langle N \rangle$  of water molecules in a volume  $\Delta V$  inside the nanotube is determined by the difference of the local excess chemical potential  $\mu_{\text{in}}^{\text{ex}}$  relative to that of the bulk fluid,  $\mu_{\text{w}}^{\text{ex}}$ :

$$\langle N \rangle = \rho_0 \Delta V \exp[-\beta(\mu_{\text{in}}^{\text{ex}} - \mu_{\text{w}}^{\text{ex}})] \quad (1)$$

where  $\mu_{\text{in}}^{\text{ex}} - \mu_{\text{w}}^{\text{ex}}$  corresponds to a potential-of-mean-force difference for water;  $\rho_0$  is the bulk water density; and  $\beta^{-1} = k_B T$ , with  $k_B$  Boltzmann’s constant, and  $T$  the temperature.

Excess chemical potentials  $\mu^{\text{ex}}$  are directly related to the distributions  $p_{\text{bind}}(u)$  of binding energies of individual molecules:

$$\exp(\beta \mu^{\text{ex}}) = \langle \exp(\beta u) \rangle = \int p_{\text{bind}}(u) \exp(\beta u) du \quad (2)$$

The binding energy  $u$  of a given water molecule is the potential-energy difference of the system in a given configuration with and without that molecule. The distribution  $p_{\text{bind}}(u)$  of binding energies is related to the distribution  $p_{\text{ins}}(u)$  of potential energies  $u$  of water molecules randomly inserted into the volume  $\Delta V$ , and averaged over equilibrium configurations of the unperturbed system<sup>27</sup>:

$$\frac{p_{\text{bind}}(u)}{p_{\text{ins}}(u)} = \exp[\beta(\mu^{\text{ex}} - u)] \quad (3)$$

Histograms for energies of water molecule removal,  $p_{\text{bind}}(u)$ , and insertion,  $p_{\text{ins}}(u)$ , were determined for bulk water, and for the central channel of the nanotube. For the nanotube histograms  $p_{\text{bind}}(u)$  and  $p_{\text{ins}}(u)$ , water molecules were, respectively, removed from and randomly inserted into a cylindrical volume  $\Delta V$  concentric with the nanotube. The radius and height of  $\Delta V$  are 0.8 and 13.4 Å, as defined by the carbon atoms at the rim. Excess chemical potentials of water in the nanotube channel and in bulk TIP3P water were then determined from a histogram analysis<sup>27</sup> of the energy distributions by fitting  $\ln[p_{\text{bind}}(u)/p_{\text{ins}}(u)]$  to  $\beta(\mu^{\text{ex}} - u)$ .

Received 12 February; accepted 20 September 2001.

1. Gelb, L. D., Gubbins, K. E., Radhakrishnan, R. & Sliwinski-Bartkowiak, M. Phase separation in confined systems. *Rep. Prog. Phys.* **62**, 1573–1659 (1999).
2. Lum, K., Chandler, D. & Weeks, J. D. Hydrophobicity at small and large length scales. *J. Phys. Chem. B* **103**, 4570–4577 (1999).
3. Wallqvist, A. & Berne, B. J. Computer simulation of hydrophobic hydration forces on stacked plates at short range. *J. Phys. Chem.* **99**, 2893–2899 (1995).
4. Lum, K. & Luzar, A. Pathway to surface-induced phase transition of a confined fluid. *Phys. Rev. E* **56**, R6283–R6286 (1997).
5. Bolhuis, P. G. & Chandler, D. Transition path sampling of cavitation between molecular scale solvophobic surfaces. *J. Chem. Phys.* **113**, 8154–8160 (2000).
6. Stillinger, F. H. Structure in aqueous solutions of nonpolar solutes from the standpoint of scaled-particle theory. *J. Solut. Chem.* **2**, 141–158 (1973).
7. MacKay, D. H. J. & Wilson, K. R. Possible allosteric significance of water structures in proteins. *J. Biomol. Struct. Dyn.* **4**, 491–500 (1986).
8. Lynden-Bell, R. M. & Rasaiah, J. C. Mobility and solvation of ions in channels. *J. Chem. Phys.* **105**, 9266–9280 (1996).
9. Sansom, M. S. P., Shrivastava, I. H., Ranatunga, K. M. & Smith, G. R. Simulations of ion channels—watching ions and water move. *Trends Biochem. Sci.* **25**, 368–374 (2000).
10. Iijima, S. Helical microtubules of graphitic carbon. *Nature* **354**, 56–58 (1991).
11. Ajayan, P. M. & Iijima, S. Smallest carbon nanotube. *Nature* **358**, 23 (1992).
12. Liu, J. *et al.* Fullerene pipes. *Science* **280**, 1253–1256 (1998).
13. Zahab, A., Spina, L., Poncharal, P. & Marlié, C. Water-vapor effect on the electrical conductivity of a single walled carbon nanotube mat. *Phys. Rev. B* **62**, 10000–10003 (2000).
14. Wilson, M. & Madden, P. A. Growth of ionic crystals in carbon nanotubes. *J. Am. Chem. Soc.* **123**, 2101–2102 (2001).
15. Luzar, A. & Chandler, D. Hydrogen-bond kinetics in liquid water. *Nature* **379**, 55–57 (1996).
16. Pomès, R. & Roux, B. Free energy profiles for H<sup>+</sup> conduction along hydrogen-bonded chains of water molecules. *Biophys. J.* **75**, 33–40 (1998).
17. Zeidel, M. L., Ambudkar, S. V., Smith, B. L. & Agre, P. Reconstitution of functional water channels in liposomes containing purified red-cell CHIP28 protein. *Biochemistry* **31**, 7436–7440 (1992).
18. Sakmann, B., Patlak, J. & Neher, E. Single acetylcholine-activated channels show burst kinetics in presence of desensitizing concentrations of agonist. *Nature* **286**, 71–73 (1980).
19. Zhong, Q. F., Jiang, Q., Moore, P. B., News, D. M. & Klein, M. L. Molecular dynamics simulation of a synthetic ion channel. *Biophys. J.* **74**, 3–10 (1998).
20. Walqvist, A., Gallicchio, E. & Levy, R. M. A model for studying drying at hydrophobic interfaces: Structural and thermodynamic properties. *J. Phys. Chem. B* **105**, 6745–6753 (2001).
21. Beckstein, O., Biggin, P. C. & Sansom, M. S. P. A hydrophobic gating mechanism for nanopores. *J. Phys. Chem. B* (in the press).
22. Wikström, M. Proton translocation by bacteriorhodopsin and heme-copper oxidases. *Curr. Opin. Struct. Biol.* **8**, 480–488 (1998).
23. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
24. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., Nola, A. D. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
25. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
26. Cornell, W. D. *et al.* A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179–5197 (1995).
27. Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* **22**, 245–268 (1976).
28. Hummer, G., Garde, S., García, A. E., Pohorille, A. & Pratt, L. R. An information theory model of hydrophobic interactions. *Proc. Natl Acad. Sci. USA* **93**, 8951–8955 (1996).

### Acknowledgements

G.H. thanks S. Garde, L. R. Pratt, A. E. García and A. Szabo for discussions. J.C.R. and J.P.N. were supported by the National Science Foundation.

Correspondence and requests for materials should be addressed to G.H. (e-mail: hummer@helix.nih.gov).

# Equilibrating metal-oxide cluster ensembles for oxidation reactions using oxygen in water

Ira A. Weinstock\*, Elena M.G. Barbuzzi†, Michael W. Wemple†, Jennifer J. Cowan†, Richard S. Reiner‡, Dan M. Sonnen‡, Robert A. Heintz‡, James S. Bond\* & Craig L. Hill†

\* Chemistry and Pulping Research Work Unit, USDA Forest Service, Forest Products Laboratory, Madison, Wisconsin 53705, USA

† Department of Chemistry, Emory University, Atlanta, Georgia 30322, USA

‡ Department of Chemical Engineering, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA

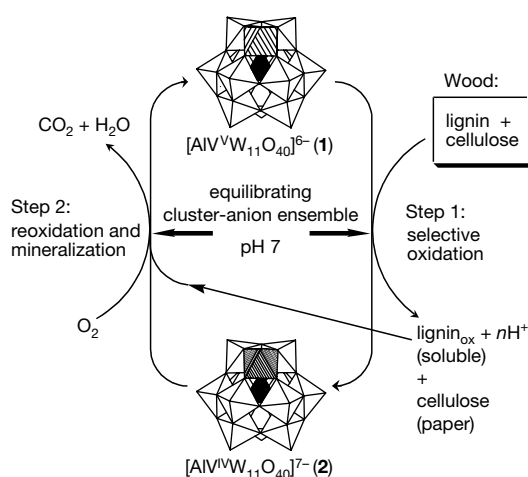
Although many enzymes can readily and selectively use oxygen in water—the most familiar and attractive of all oxidants and solvents, respectively—the design of synthetic catalysts for selective water-based oxidation processes utilizing molecular oxygen<sup>1–4</sup> remains a daunting task<sup>5,6</sup>. Particularly problematic is the fact that oxidation of substrates by O<sub>2</sub> involves radical chemistry, which is intrinsically non-selective and difficult to control. In addition, metallo-organic catalysts are inherently susceptible to degradation<sup>5</sup> by oxygen-based radicals, while their transition-metal-ion active sites often react with water to give insoluble, and thus inactive, oxides or hydroxides<sup>7</sup>. Furthermore, pH control is often required to avoid acid or base degradation of organic substrates or products. Unlike metallo-organic catalysts, polyoxometalate anions are oxidatively stable and are reversible oxidants<sup>8,9</sup> for use with O<sub>2</sub> (refs 8–10). Here we show how thermodynamically controlled self-assembly of an equilibrated ensemble of polyoxometalates, with the heteropolytungstate anion<sup>11,12</sup> [AlV<sup>V</sup>W<sub>11</sub>O<sub>40</sub>]<sup>6–</sup> as its main component, imparts both stability in water and internal pH-management. Designed to operate at near-neutral pH, this system facilitates a two-step O<sub>2</sub>-based process for the selective delignification of wood (lignocellulose)

fibres. By directly monitoring the central Al atom, we show that equilibration reactions typical of polyoxometalate anions<sup>13,14</sup> keep the pH of the system near 7 during both process steps.

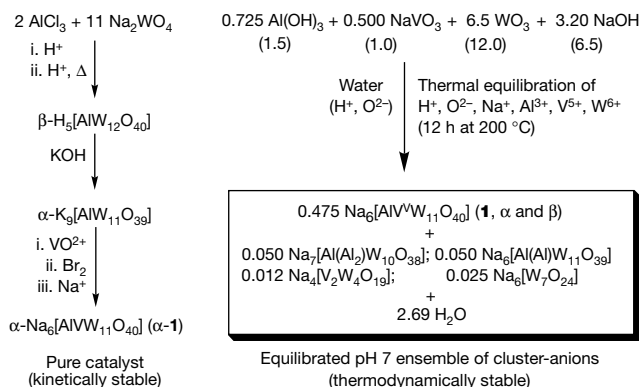
An effluent-free process using O<sub>2</sub> and H<sub>2</sub>O would be an attractive alternative to ClO<sub>2</sub>-based pulp bleaching<sup>15</sup>, which is globally one of the largest industrial oxidation processes and generates considerable waste. To explore fundamental questions concerning the suitability of polyoxometalate (POM) ions for delignification processes, we use [AlV<sup>V</sup>W<sub>11</sub>O<sub>40</sub>]<sup>6–</sup> (**1** in Fig. 1), which can be spectroscopically monitored in solution.

Like many<sup>8</sup> POM ions, **1** is a reversible oxidant, capable of undergoing repeated cycles of reduction and re-oxidation<sup>9</sup>. The electrochemical potential for the one-electron reduction of **1** to [AlV<sup>IV</sup>W<sub>11</sub>O<sub>40</sub>]<sup>7–</sup> (**2** in Fig. 1) at pH 7 in water (0.480 V versus the normal hydrogen electrode, NHE) lies below that of the O<sub>2</sub>/H<sub>2</sub>O couple (0.815 V versus NHE). Thus, whereas electron transfer from a variety of inorganic or organic substrates (such as lignin) to **1** can be highly selective (that is, the relative rates of these oxidation reactions span a wide range of values), electron transfer from **2** to O<sub>2</sub> (reoxidation of **2** to **1**) proceeds spontaneously<sup>10</sup>. These steps—anaerobic reduction of **1** by substrate and reoxidation of **2** by O<sub>2</sub>—sum to electron transfer from substrate to O<sub>2</sub>. Further, the reduction of O<sub>2</sub> by reduced heteropolytungstate anions ('heteropolyblues') such as **2** initially gives O<sub>2</sub><sup>•–</sup>, and subsequently and rapidly H<sub>2</sub>O<sub>2</sub> (refs 10, 16, 17). By making use of the inevitable and highly reactive intermediates of radical-chain oxidation (for example, O<sub>2</sub><sup>•–</sup>, HO<sup>•</sup>, their organic analogues and others)<sup>18</sup>, the reoxidation of **2** by O<sub>2</sub> can be carried out under conditions (time, temperature and O<sub>2</sub> pressure) that oxidatively eliminate (mineralize) the inevitable by-products of substrate oxidation<sup>19</sup>. In the effluent-free process shown in Fig. 1, solutions of **1** selectively depolymerize lignin in wood fibres by direct oxidation (step 1, reduction of **1** to **2** by lignin; no O<sub>2</sub> present) and subsequently catalyse the O<sub>2</sub>-oxidation of solubilized lignin fragments to their biosynthetic precursors CO<sub>2</sub> and H<sub>2</sub>O (step 2).

Development of the equilibrated solutions is summarized in Fig. 2. In order to optimize and fully characterize all components of the desired equilibrated solution, isomerically pure samples of



**Figure 1** Proposed two-step process for conversion of lignin in wood to CO<sub>2</sub> and H<sub>2</sub>O. (Wood is a composite of lignin and cellulose<sup>15</sup>.) The structures of polyoxometalates (POMs) α-[AlV<sup>V</sup>W<sub>11</sub>O<sub>40</sub>]<sup>6–</sup> and α-[AlV<sup>IV</sup>W<sub>11</sub>O<sub>40</sub>]<sup>7–</sup> (α-**1** and α-**2**) are shown in polyhedral notation. The central, tetrahedrally coordinated Al atoms are shown in black. Step 1 (right) is the selective anaerobic oxidation of substrate (lignin) and reduction of POM (**1**→**2**). Step 2 is the reoxidation of POM (**2**→**1**) by O<sub>2</sub> with simultaneous POM-catalysed oxidation of the lignin fragments to CO<sub>2</sub> and H<sub>2</sub>O (mineralization). Reversible acid-condensation reactions of cluster-anions (POM species) in equilibrium with **1** and **2** maintain the pH near neutral.

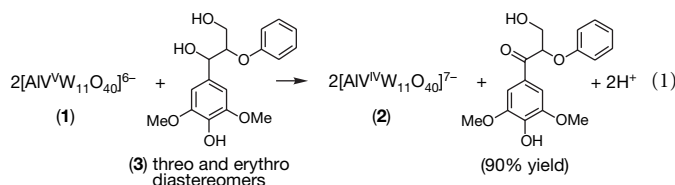


**Figure 2** Stepwise synthesis of kinetically stable pure **1**, and equilibration to give a thermodynamically stable reaction medium. Three attributes of the reaction medium (right), pertinent to transition-metal-mediated transformations in water, follow: (1) it is uniquely stable with respect to precipitation of redox-active metal ions (all elements present, H, Na, Al, V, W and O are in thermal equilibrium with one another), (2) it is stable to oxidative degradation (all main-group and transition-metal cations present, H<sup>+</sup>, Na<sup>+</sup>, Al(III) and the d<sup>0</sup> cations V(V) and W(VI), are in their highest accessible valence states), and (3) the cluster anions necessary for internal H<sup>+</sup> management are themselves integral components of the thermodynamically controlled (equilibrating) reaction medium. As such, self-buffering is an inherent and reversible property of the electron transfer and catalyst system as a whole, attributable to changes in the position of equilibrium of the entire molecular ensemble.

the target POM catalyst, in this case  $\alpha\text{-Na}_6[\text{AlVW}_{11}\text{O}_{40}]$  ( $\alpha\text{-Na}_6\text{I}^{12}$ ), as well as precursors<sup>20</sup> and closely related derivatives must first be isolated. This is accomplished using the method by which POMs have traditionally been prepared<sup>11,12</sup>—sequential, kinetically controlled steps of acid-condensation and base-hydrolysis (Fig. 2, left).

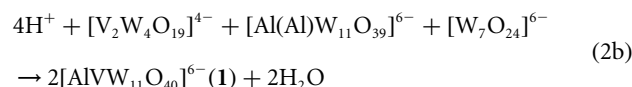
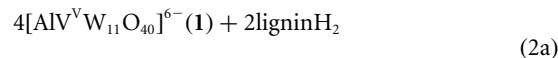
Next, **1** is prepared in one step using a thermodynamically controlled process: the thermal equilibration of component main-group and transition-metal oxides and hydroxides in water (Fig. 2, right). The final position of equilibrium, which dictates the pH and self-buffering properties of the medium, is controlled by adjusting the relative concentrations of the component precursors. Use of the molar concentrations shown as prefixes in Fig. 2 (stoichiometric ratios are in parentheses) gives a thermally equilibrated pH 7 solution of **1** ( $90 \pm 5\%$  relative to  $\text{NaVO}_3$ ; mixture of  $\alpha$  and  $\beta$  isomers) as its major constituent. Final solution concentrations were determined to within  $\pm 5\%$  by  $^{27}\text{Al}$ ,  $^{51}\text{V}$  and  $^{183}\text{W}$  NMR spectroscopy<sup>12,21</sup>; data concerning  $\text{Na}_7[\text{Al}(\text{Al}_2)\text{W}_{10}\text{O}_{38}]$  are available as Supplementary Information. The cluster anions in equilibrium with **1** are shown in the box at the bottom right in Fig. 2. Any  $\text{H}^+$  or  $\text{OH}^-$  ions introduced during use (Fig. 1; that is, reversible reduction of **1** by substrate or subsequent reoxidation by  $\text{O}_2$ ) are consumed, respectively, by reversible acid-condensation or base-hydrolysis reactions which are typical of these cluster-anions<sup>22</sup>. Analogous reactions have been quantified in fundamental studies of cation speciation and cluster-anion formation<sup>13</sup>. In the reaction system in Fig. 1, these physical properties are incorporated by design to provide inherent stability and internal  $\text{H}^+$  management during electron transfer<sup>10</sup> and catalysis. Thus, the position of equilibrium of the reaction medium changes in response to the addition of  $\text{H}^+$  simultaneous with the addition of electrons (Fig. 1, step 1; electron transfer from substrate to **1**), and to the addition of “ $\text{O}^{2-}$  equivalents” simultaneous with the removal of electrons (Fig. 1, step 2; electron transfer from **2** to  $\text{O}_2$ )<sup>14</sup>.

In a typical laboratory-scale delignification reaction, unbleached softwood pulp from red pine and jack pine (5% lignin by weight) was heated in an equilibrated pH 7 solution of **1** (prepared as in Fig. 2; 0.01 g of pulp per ml of 0.475 M **1**) at 130 °C for three hours under argon. During the reaction, 75% of the lignin<sup>23</sup> present in the unbleached pulp fibres is oxidatively depolymerized to water-soluble fragments, while  $\sim 15\%$  of cluster **1** is reduced to **2**. The oxidation of lignin or other organic compounds in  $\text{H}_2\text{O}$  releases  $\text{H}^+$  ions, as illustrated by equation (1), which shows the partial oxidation by **1** (12 h at 22 °C) of a dimeric model<sup>24</sup> (**3**) for polymeric lignin:



Although the reaction in equation (1) demonstrates the need for pH management, **3** is more readily oxidized, and thus represents a less stringent delignification test than the non-phenolic lignin models, such as veratryl alcohol derivatives, that are commonly used in pulp bleaching studies. Moreover, POM delignification requires cleavage of the lignin polymer, and by inference, of **3** as well. Room-temperature cleavage of **3** by the POM  $\alpha\text{-Na}_5[\text{SiVW}_{11}\text{O}_{40}]$ , a close analogue of **1**, occurs via two sequential one-electron oxidation steps to give the corresponding cyclohexadienyl cation, followed by hydrolysis to give monomeric 2,6-dimethoxy-*p*-benzoquinone and other products<sup>24</sup>. Reaction of **3** with **1** at higher temperatures gives mixtures of products, including 2,6-dimethoxy-*p*-benzoquinone.

As is evident from equation (1), the number of  $\text{H}^+$  ions released during lignin oxidation and cleavage in water corresponds to the number of equivalents of **1** reduced to **2**. Thus, during delignification of the unbleached fibres, 0.07 moles per litre of  $\text{H}^+$ —15% conversion of **1** (0.475 M) to **2**—are generated (equation (2a)). Meanwhile, the pH of the solution remains unchanged as  $\text{H}^+$  ions are converted to water by reaction with the oxo ligands of cluster-anions present in the equilibrium mixture (equation (2b)), a process quantified by  $^{27}\text{Al}$ ,  $^{51}\text{V}$  and  $^{183}\text{W}$  NMR spectroscopy:



(Note that the target catalyst, **1**, is itself a product.) To ensure reversibility, precipitation of fully condensed tungsten oxide (that is, solid  $\text{WO}_3$ ) must be prevented. This is accomplished by preferential co-condensation of  $\text{W}(\text{vi})$  with  $\text{Al}(\text{iii})$  to give soluble tungstoaluminate cluster-anions<sup>25</sup> (for example, **1** in equation (2b)). Hence, at greater percentages of reduction, the more Al-rich anion,  $\text{Na}_7[(\text{Al})(\text{Al}_2)\text{W}_{10}\text{O}_{38}]$  (Fig. 2; effectively a reservoir for Al), reacts as well. If not for these acid-induced, POM self-assembly reactions, the pH of the solution would decrease to 1.15. At this low pH value, the  $\beta$ -D-glycosidic linkages between glucose units in cellulose would be readily cleaved and thus, the desired product would be degraded.

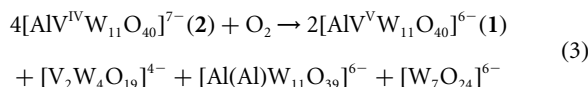
The use of **1** and other POMs ensures highly selective oxidation of lignin alone, whereas bleaching of wood pulp by  $\text{O}_2$  has been considered not useful for delignification to low percentages of lignin<sup>26</sup>, because it results in the oxidative degradation of cellulose fibres by non-selective radical-chain autooxidation reactions. Degradation of cellulose fibres is typically indicated by a decrease in pulp viscosity<sup>23</sup>—this viscosity is a nonlinear function of the average chain-length of the cellulose polymer. After reaction of unbleached fibres (initial viscosity 31 mPa s) with **1** (0.01 g of pulp per ml of 0.475 M equilibrated **1** for 50 min at 145 °C) followed by filtering, washing, and mild extraction with dilute aqueous NaOH, 81% of the lignin is removed. The final viscosity is 22 mPa s. These values are comparable to those typically obtained industrially using chlorine compounds and far exceed those obtained using  $\text{O}_2$  directly (final viscosities of less than 15 mPa s).

If systems such as **1** are to be used commercially, one issue that must be addressed is the need for near-quantitative recovery of cluster anions from delignified wood fibres, given that the cluster-anions mediate the delignification reactions stoichiometrically, rather than in a truly catalytic fashion. Keggin anions (close-packed structures) possess hydrodynamic radii in water of  $\sim 5.6$  Å (ref. 27), similar in size to typical lignin monomers (ring-methoxylated and side-chain hydroxylated phenylpropane derivatives) and small enough to diffuse freely between cellulose fibres within the wood-cell wall. In addition, due to the organic acids present on their surfaces, the polysaccharide polymers (cellulose and hemicelluloses) present in wood fibres possess a net negative charge at near-neutral pH values. Hence, as a consequence of both size and charge, the POM clusters are easily washed from fibre samples at diffusion-controlled rates. In preliminary washing studies (see Supplementary Information), 0.12 p.p.m. V and 4.0 p.p.m. W were detected in fibres after washing and alkali extraction (a standard industry practice). These values are comparable to naturally occurring levels in the environment<sup>28,29</sup> and in selected foods<sup>28,29</sup>.

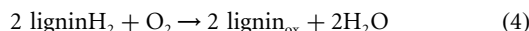
After removing the delignified cellulose fibres, reoxidation of **2** to **1** by  $\text{O}_2$  is carried out in a separate step (step 2 in Fig. 1). The four-electron reduction of  $\text{O}_2$  in  $\text{H}_2\text{O}$  results in the production of 4 equivalents of  $\text{OH}^-$  (effectively the reaction of two “ $\text{O}^{2-}$ —equivalents” with 2  $\text{H}_2\text{O}$ ). Each equivalent of **2** reoxidized to **1** thus introduces a



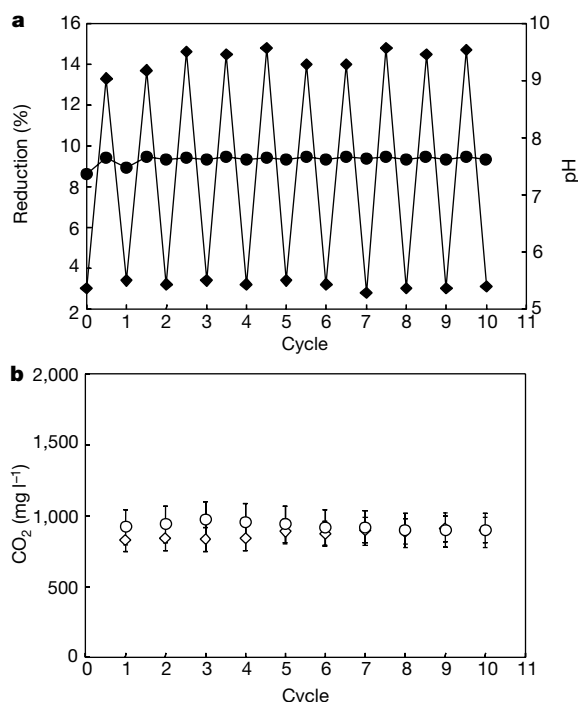
like quantity of  $\text{OH}^-$ . In the absence of a mechanism for managing pH, reduction of  $\text{O}_2$  by the pH-neutral 0.07 M solution of **2** (and the production of 0.07 M  $\text{HO}^-$ ) would thus result in an increase in pH to 12.8. This increase is prevented, however, by rapid base-hydrolysis of the cluster-anions formed during delignification. These hydrolysis reactions, effectively the reaction of " $\text{O}^{2-}$ -equivalents" from reduction of  $\text{O}_2$  with previously condensed cluster-anions, return the system to its initial position of equilibrium:



Equations (2) and (3) sum to equation (4), the selective oxidation of lignin (lignin $\text{H}_2$ ) by  $\text{O}_2$  to give water-soluble lignin fragments (lignin $_{\text{ox}}$ ) and  $\text{H}_2\text{O}$ :



Moreover, when the reaction shown in equation (3) is carried out at 210 °C for 3 h under  $\text{O}_2$  (2 MPa), **1** catalyses the conversion of solubilized lignin fragments (lignin $_{\text{ox}}$  in equations (2a) and (4)) to  $\text{CO}_2$  and  $\text{H}_2\text{O}$  (oxidative mineralization). The amount of  $\text{CO}_2$  evolved ( $930 \pm 40 \text{ mg CO}_2 \text{ per l}$ ) corresponds to the mass (lignin and some polysaccharides) earlier removed from the pulp fibres (see ref. 19 and work cited therein). These reactions (equations (2–4) and mineralization) define a process currently achieved only in nature: the complete and selective  $\text{O}_2$ -oxidation of lignin to  $\text{CO}_2$  and  $\text{H}_2\text{O}$ .

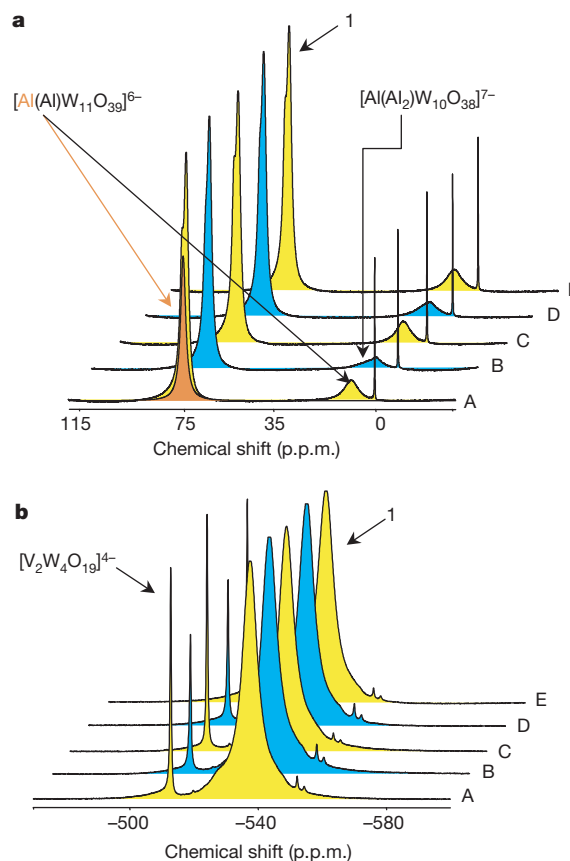


**Figure 3** Oxidation of lignin in wood-pulp fibres to  $\text{CO}_2$  and  $\text{H}_2\text{O}$ . Details are given in Methods. A single equilibrated solution of **1** prepared as shown in Fig. 2 was used in 10 successive two-step cycles of delignification (anaerobic oxidation by **1**) and subsequent  $\text{O}_2$ -reoxidation/lignin mineralization. In the first step of each cycle, unbleached pulp was heated in a solution of **1**. Lignin is oxidatively cleaved and solubilized by **1**, which is thereby reduced. The delignified cellulose fibres were then removed and  $\text{O}_2$  was used (second step of each cycle) to reoxidize **2** to **1** and to convert solubilized lignin fragments to  $\text{CO}_2$  and  $\text{H}_2\text{O}$ . **a**, Percentage reduction of **1** to **2** (diamonds, left scale) and solution pH values (circles, right scale). Reversible acid-condensation (self-buffering) reactions keep the pH near neutral throughout all 10 cycles. **b**, Measured (circles) and theoretical (diamonds) values of  $\text{CO}_2$  produced during the catalytic mineralization of solubilized lignin fragments.

As a model for continuous operation, the equilibrated solution was used in nine additional cycles of delignification and reoxidation/mineralization (Fig. 3). Notably, the pH of the POM solution remained nearly constant throughout the ten cycles (Fig. 3a, right scale). The  $\text{CO}_2$  produced in each mineralization step was comparable to the amount expected on the basis of the masses of lignin removed and yields of the delignified pulp samples (Fig. 3b; see Methods and Supplementary Information).

Four indices plotted in Fig. 3, all of which remained constant, provide evidence that the integrity of the catalyst system was retained throughout the multi-cycle experiment: (1) the amount of **1** reduced during delignification; (2) the pH of the solution; (3) the amount of **2** reoxidized by  $\text{O}_2$ ; and (4) the amount of  $\text{CO}_2$  produced. In addition, the amount of lignin removed in each cycle remained constant ( $75 \pm 5\%$ ). More direct observation of the reversible behaviour and stability of the equilibrating system is provided by  $^{27}\text{Al}$ ,  $^{51}\text{V}$  and  $^{183}\text{W}$  NMR analysis of the solution before, during and after the ten-cycle experiment ( $^{27}\text{Al}$  and  $^{51}\text{V}$  NMR spectra are shown in Fig. 4).

Although the spectroscopically accessible prototype system is



**Figure 4** NMR analysis of the solution used in the 10-cycle experiment in Fig. 3. Samples were analysed by high-field  $^{27}\text{Al}$  and  $^{51}\text{V}$  NMR spectroscopy: **a**,  $^{27}\text{Al}$  NMR spectra of the freshly prepared solution (spectrum A), after reduction by lignin in cycle 1 (reoxidized by  $\text{Br}_2$  at room temperature) (B), after  $\text{O}_2$ -reoxidation/mineralization in cycle 1 (C), after reduction by lignin in cycle 10 (D) and after final  $\text{O}_2$ -reoxidation/mineralization in cycle 10 (E). **b**,  $^{51}\text{V}$  NMR spectra of the freshly prepared solution (spectrum A), after reduction in cycle 1 (reoxidized by  $\text{Br}_2$  at room temperature) (B), after reoxidation in cycle 1 (C), after reduction in cycle 10 (D) and after final reoxidation in cycle 10 (E). See Methods for assignments. Integration of  $^{27}\text{Al}$  and  $^{51}\text{V}$  NMR signal intensities and spectral deconvolution (an example is shown in orange in  $^{27}\text{Al}$  NMR spectrum A) were used to quantify the reactions in equations (2b) and (3). Signals due to cluster-anions present in freshly prepared or  $\text{O}_2$ -oxidized solutions are shown in yellow, while those due to cluster-anions in solutions reduced by lignin (subsequently reacted with  $\text{Br}_2$  at room temperature) are shown in blue.

ideal for characterizing the chemical changes (Fig. 4 and equations 2 and 3) involved, any practically useful POM ensemble will need to be able to delignify more concentrated pulp mixtures (that is, 80–120 g pulp per l) with greater efficiency (defined as lignin solubilized per mole of POM clusters; Fig. 1, step 1). Moreover, practical implementation will require milder operation conditions and faster turnover rates ( $T < 250^\circ\text{C}$  and  $t < 1\text{ h}$ ) in wet oxidation (Fig. 1, step 2), and near-quantitative yet efficient and economic POM recovery from the treated pulp in industrial settings.

Systems developed further to meet some of these demanding criteria for practicality include a pH 5.5 equilibrating ensemble, designated “ $\text{Na}_{6.9}\text{SiV}_{0.9}\text{Mo}_{1.0}\text{W}_{10.1}$ ” in which the dominant species are Keggin anions, structurally analogous to **1**, of formula  $[\text{SiVMoW}_{10}\text{O}_{40}]^{5-}$  and/or  $[\text{SiVW}_{11}\text{O}_{40}]^{5-}$  (see Methods and Supplementary Information). Unlike solutions of **1**, the optimized systems are difficult to study by NMR because they include isostructural clusters containing both W(vi) and Mo(vi), a relatively NMR-insensitive Si(IV) heteroatom and, as below, paramagnetic Mn cations. The “ $\text{Na}_{6.9}\text{SiV}_{0.9}\text{Mo}_{1.0}\text{W}_{10.1}$ ” system was used to delignify commercial fibre samples (20-litre reactions at 30 g pulp per l) supplied by three paper manufacturers (key fibre properties indicative of high selectivity are available as Supplementary Information). Typically, pH values of the reaction medium increase during delignification from 5.5 to near 6 (addition of  $\text{H}^+$ , Fig. 1, step 1; 1.5 h at  $140^\circ\text{C}$ ), and decrease to their original values during reoxidation (addition of “ $\text{O}^{2-}$  equivalents”).

More rapid rates of POM reoxidation by  $\text{O}_2$  (Fig. 1, step 2; 0.5–1.0 h,  $210^\circ\text{C}$ , 300 p.s.i.  $\text{O}_2$ ), were achieved in model reduction–oxidation cycles upon partial replacement of W(vi) by 0.05 to 0.10 molar equivalents of Mn(III) (see Methods), which is known to catalyse wet-oxidation reactions<sup>18</sup>. Upon complete (~100%) reduction of a 0.5 M solution by gaseous CO (4–8 equiv. at  $200^\circ\text{C}$ ; reduction of all V(v) present to V(IV)), the pH increased from about 5.5 to 6.8 during the effective addition of ~0.5 M  $\text{H}^+$ . Reoxidation by  $\text{O}_2$  (and the effective addition of 0.5 M  $\text{OH}^-$ ) resulted in a decrease in pH to its initial value (5.5). These values are directly transferable to delignification systems and are a good indication of expected performance in actual practice. Although 100% reduction is difficult to achieve under laboratory conditions, similar results (reoxidation rates and self buffering) are observed. The superficially counter-intuitive pH changes, previously observed in reductions of labile molybdovanadophosphate systems<sup>14,30</sup>, are due to changes in the relative  $\text{H}^+$ -dependent thermodynamic stabilities of cluster anions upon reduction and reoxidation—reduced anions are typically stable at higher pH values<sup>30</sup>.

Although industrial use of POM-based delignification systems will require further development, the data in Figs 2–4 show that the thermodynamically controlled self-assembly of inorganic oxides can yield complex, yet highly organized, supramolecular systems that behave as stable, pH-controlled and multi-functional (electron transfer and catalyst) aqueous reaction media. □

## Methods

### Multi-cycle experiment in Fig. 3

In the first step of each cycle, unbleached pulp (5% lignin by weight) was stirred under argon in a 0.475 M solution of **1** (0.01 g pulp per ml; initially 2.5 g pulp (dry weight), in 250 ml solution) in a 1.0-litre Parr high-pressure vessel and heated for three hours at  $130^\circ\text{C}$ . During each reaction, ~75% of the lignin was oxidatively cleaved and solubilized by **1**, which was thereby reduced (Fig. 3a, % reduction of **1** to **2** (diamonds), left scale; determined by UV–visible spectroscopy,  $\epsilon_{680\text{nm}} = 180\text{ cm}^{-1}\text{ M}^{-1}$ , and solution pH values (circles), right scale). The delignified cellulose fibres were then removed and washed. To avoid irreproducible losses of volatile organic compounds, the small volumes of POM solution washed from the fibres were discarded rather than concentrated and returned—as practical operation would require—to the system. Removal of analytical samples also contributed to the total decrease in volume of the POM solution to 187 ml (reacted with 1.87 g unbleached pulp) at the beginning of cycle 10. In the second step of each cycle,  $\text{O}_2$  was used to reoxidize **2** to **1** (Fig. 3a, (diamonds), left scale) and to convert solubilized lignin fragments to  $\text{CO}_2$  (Fig. 3b) and  $\text{H}_2\text{O}$  (see Supplementary Information). In each case,

the reactor was fitted with a gas-entrainment impeller, sealed under 2 MPa  $\text{O}_2$  (static pressure) and heated for 3 h at  $210^\circ\text{C}$ . Reversible cluster-anion equilibration (self-buffering) reactions keep the pH near neutral throughout all 10 cycles (Fig. 3a, (circles), right scale). Experimentally measured  $\text{CO}_2$  values (Fig. 3b, circles) were determined by reaction of head gases with aqueous  $\text{Ba}(\text{OH})_2$ ; associated uncertainties are minimum values based on titrations of  $\text{Ba}(\text{OH})_2$  solutions. Theoretical values (Fig. 3b; diamonds) were calculated by mass balance from decreases in lignin content and final pulp yields; associated uncertainties were estimated from the variation observed in replicate measurements of identical fibre samples. For lignin, calculations were based on the molecular weight of repeating  $\beta$ -O-4 linked coniferyl alcohol subunits<sup>15</sup> ( $(\text{C}_{10}\text{H}_{12}\text{O}_4)_n + 11\text{ O}_2 \rightarrow 10\text{ CO}_2 + 6\text{ H}_2\text{O}$ ).  $\text{CO}_2$  from any additional decreases in pulp mass was estimated using repeating glucose subunits,  $(\text{C}_6\text{H}_{10}\text{O}_5)_n$ , to represent cell-wall polysaccharides.

### NMR data in Fig. 4

Samples of the solution used in the 10-cycle experiment in Fig. 3 were analysed by high-field  $^{27}\text{Al}$  and  $^{51}\text{V}$  NMR spectroscopy (paramagnetic species were oxidized by  $\text{Br}_2$ ). All  $^{27}\text{Al}$  NMR spectra are externally referenced to 0.1 M aqueous  $\text{AlCl}_3$  ( $[\text{Al}(\text{H}_2\text{O})_6]^{3+}$ ,  $\delta = 0\text{ p.p.m.}$ ). The signals at 70–80 p.p.m. arise from tetrahedrally coordinated Al atoms located at the centre of the identified cluster-anions<sup>12</sup>; signals at 8–15 p.p.m. arise from octahedrally coordinated Al atoms. Integration and spectral deconvolution (an example is shown in red in Fig. 4a spectrum A) was used to quantify reversible changes in the position of equilibrium upon successive additions of  $\text{H}^+$  (delignification, equation (2)) and of “ $\text{O}^{2-}$ -equivalents” ( $\text{O}_2$ -reoxidation, equation (3)). All  $^{51}\text{V}$  NMR spectra are referenced to pure  $\text{VOCl}_3$ . Consumption and re-formation of  $[\text{V}_2\text{W}_2\text{O}_{19}]^{4-}$  (equations (2b) and (3)) is shown by changes in intensity of the sharp signal at –511 p.p.m. The large linewidth of the signal at –535 p.p.m. (from  $^{51}\text{V}$  in **1**) is due to quadrupolar broadening by  $^{27}\text{Al}$  ( $I = 5/2$ )<sup>21</sup> and to the presence of an equilibrium mixture of  $\alpha$ - and  $\beta$ -Keggin structural isomers<sup>12</sup>.

### Preparation of “ $\text{Na}_{6.9}\text{SiV}_{0.9}\text{MoW}_{10.1}$ ”

To a Parr reactor containing ~9.90 litre water was added 1.0 equiv. (993.1 g)  $\text{Na}_2\text{SiO}_3$  (anhydrous, 92.2%), 4.9 equiv. (1,500.0 g)  $\text{NaOH}$  (98%), 0.45 equiv. (618.8 g)  $\text{V}_2\text{O}_5$  (99.2%), 1.0 equiv. (1,090.5 g)  $\text{MoO}_3$  (99%) and 10.1 equiv. (17,583.7 g)  $\text{WO}_3$  (99.88%). The reactor vessel was sealed and heated to  $200^\circ\text{C}$  for 3 hours under 100 p.s.i. (cold) of  $\text{O}_2$  with continuous rapid stirring (about 500 r.p.m.). Before use in delignification, the solution was treated with ozone in a glass vessel and concentrated to about 0.6 M (density 2.422). When combined with pulp for use in delignification, water in the pulp fibres contributes to give a final concentration of 0.5 M relative to Si(IV).

### Preparation of “ $\text{Na}_{6.5}\text{SiV}_{0.9}\text{Mn}_{0.1}\text{MoW}_{10}$ ”

To a Parr reactor containing ~188 ml water was added 1.0 equiv.  $\text{Na}_2\text{SiO}_3$ , 4.5 equiv.  $\text{NaOH}$ , 0.45 equiv.  $\text{V}_2\text{O}_5$ , 1.0 equiv.  $\text{MoO}_3$ , 10.0 equiv.  $\text{WO}_3$  and 0.1 equiv.  $\text{MnO}_2$  (during synthesis, Mn(IV) is reduced, most probably by water, to Mn(III)). Oxygen (200–250 p.s.i.) was added and the mixture stirred (1,000 r.p.m.) while heated to  $200^\circ\text{C}$  for 5 h, then cooled, vented, opened and the solution filtered. Solutions possess densities of about  $2.2\text{ g ml}^{-1}$  and have pH values in the range 5–6.

Received 14 February; accepted 18 September 2001.

- Groves, J. T. & Quinn, R. Aerobic epoxidation of olefins with ruthenium porphyrin catalysts. *J. Am. Chem. Soc.* **107**, 5790–5792 (1985).
- Neumann, R. & Dahan, M. Molecular oxygen activation: A ruthenium substituted polyoxometalate as an inorganic dioxygenase. *Nature* **388**, 353–355 (1997).
- Weiner, H. & Fink, R. G. An all-inorganic, polyoxometalate-based catechol dioxygenase that exhibits >100,000 catalytic turnovers. *J. Am. Chem. Soc.* **121**, 9831–9842 (1999).
- Döbler, C., Mehlretter, G. M., Sundermeier, U. & Beller, M. Osmium-catalyzed dihydroxylation of olefins using dioxygen or air as the terminal oxidant. *J. Am. Chem. Soc.* **122**, 10289–10297 (2000).
- Hill, C. L. & Weinstock, I. A. Homogeneous catalysis: On the trail of dioxygen activation. *Nature* **388**, 332–333 (1997).
- Hill, C. L. Homogeneous catalysis: Controlled green oxidation. *Nature* **401**, 436–437 (1999).
- Baer, C. F. Jr & Mesmer, R. E. *The Hydrolysis of Cations* (Wiley & Sons, New York, 1976).
- Hill, C. L. Special thematic issue on polyoxometalates. *Chem. Rev.* **98**, 1–390 (1998).
- Neumann, R. Polyoxometalate complexes in organic oxidation chemistry. *Prog. Inorg. Chem.* **47**, 317–370 (1998).
- Weinstock, I. A. Homogeneous-phase electron-transfer reactions of polyoxometalates. *Chem. Rev.* **98**, 113–170 (1998).
- Pope, M. T. & Müller, A. Polyoxometalate chemistry: An old field with new dimensions in several disciplines. *Angew. Chem. Int. Edn Engl.* **30**, 34–48 (1991).
- Weinstock, I. A., Cowan, J. J., Barbuzzi, E. M. G., Zeng, H. & Hill, C. L. Equilibria between  $\alpha$  and  $\beta$  isomers of kegglin heteropolytungstates. *J. Am. Chem. Soc.* **121**, 4608–4617 (1999).
- Andersson, L., Hastings, J. J., Howarth, O. W. & Pettersson, L. Aqueous tungstovanadate equilibria. *J. Chem. Soc. Dalton Trans.* 2705–2711 (1996).
- Grate, J. H., Hamm, D. R. & Mahajan, S. In *Polyoxometalates: From Platonic Solids to Anti-Retroviral Activity* (eds Pope, M. T. & Müller, A.) 281–306 (Kluwer, Dordrecht, 1994).
- Sjöström, E. *Wood Chemistry: Fundamentals and Applications* (Academic, San Diego, 1992).
- Hiskia, A. & Papaconstantinou, E. Photocatalytic oxidation of organic compounds by polyoxometalates of molybdenum and tungsten. Catalyst regeneration by dioxygen. *Inorg. Chem.* **31**, 163–167 (1992).
- Duncan, D. C. & Hill, C. L. Mechanism of reaction of reduced polyoxometalates with  $\text{O}_2$  evaluated by  $^{17}\text{O}$  NMR. *J. Am. Chem. Soc.* **119**, 243–244 (1997).
- Birchmeier, M. J., Hill, C. G. J., Houtman, C. J., Atalla, R. H. & Weinstock, I. A. Enhanced wet air oxidation: synergistic rate acceleration upon effluent recirculation. *Ind. Eng. Chem. Res.* **39**, 55–64 (2000).

19. Weinstock, I. A., Atalla, R. H., Reiner, R. S., Houtman, C. J. & Hill, C. L. Selective transition-metal catalysis of oxygen delignification using water-soluble salts of polyoxometalate (POM) anions. Part I. Chemical principles and process concepts. *Holzforschung* **52**, 304–310 (1998).
20. Akitt, J. W. & Farthing, A. Aluminium-27 nuclear magnetic resonance studies of heteropolyanions containing aluminium as heteroatom. *J. Chem. Soc. Dalton Trans.* 1615–1616 (1981).
21. Akitt, J. W. in *Modern NMR Techniques and their Applications in Chemistry* (eds Popov, A. I. & Hallenga, K.) Ch. 7, 427–483 (Dekker, New York, 1991).
22. Pope, M. T. *Heteropoly and Isopoly Oxometalates* (Springer, Berlin, 1983).
23. *TAPPI Test Methods 1991* (Tappi Press, Atlanta, Georgia, 1991).
24. Weinstock, I. et al. Selective transition-metal catalysis of oxygen delignification using water-soluble salts of polyoxometalate (POM) anions. Part II. Reactions of alpha-[SiVW<sub>11</sub>O<sub>40</sub>]<sup>5-</sup> with phenolic lignin-model compounds. *Holzforschung* **52**, 311–318 (1998).
25. Carrier, X., de la Caillerie, J. B. d. E., Lambert, J. F. & Che, M. The support as a chemical reagent in the preparation of WO<sub>3</sub>/gamma-Al<sub>2</sub>O<sub>3</sub> catalysts: Formation and deposition of aluminotungstic heteropolyanions. *J. Am. Chem. Soc.* **121**, 3377–3381 (1999).
26. Dence, C. W. & Reeve, D. W. (eds) *Pulp Bleaching: Principles and Practice* (Tappi Press, Atlanta, Georgia, 1996).
27. Grigorov, V. A., Hill, C. L. & Weinstock, I. A. Role of alkali metal cation size in the energy and rate of electron-transfer to solvent-separated 1:1 [(M<sup>+</sup>)(Acceptor)] (M<sup>+</sup> = Li<sup>+</sup>, Na<sup>+</sup>, K<sup>+</sup>) ion pairs. *J. Am. Chem. Soc.* **123**, 5292–5307 (2001).
28. Myron, D. R., Givand, S. H. & Nielsen, F. H. Vanadium content of selected foods as determined by flameless atomic absorption spectroscopy. *J. Agric. Food Chem.* **25**, 297–299 (1977).
29. Kletzin, A. & Adams, M. W. W. Tungsten in biological systems. *FEMS Microbiol. Rev.* **18**, 5–63 (1996).
30. Pettersson, L. in *Polyoxometalates: From Platonic Solids to Anti-Retroviral Activity* (eds Pope, M. T. & Müller, A.) 27–40 (Kluwer, Dordrecht, 1994).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

### Acknowledgements

I.A.W. is a visiting scientist at the Department of Chemistry, Emory University, but based at the Forest Products Laboratory. We thank W. E. Bowerman, H. Zeng, S. Ralph and L. L. Landucci for technical assistance; we also acknowledge the contributions of R. H. Atalla, M. J. Birchmeier, C. G. Hill Jr, C. J. Houtman, S. E. Reichel and E. L. Springer. This work was supported by the DOE, NSF and member companies of the USDA Forest Service, Polyoxometalate Bleaching Consortium.

Correspondence and requests for materials should be addressed to I.A.W. (e-mail: iaweinst@facstaff.wisc.edu) or C.L.H. (e-mail: chill@emory.edu).

## Determinants of establishment success in introduced birds

Tim M. Blackburn\* & Richard P. Duncan†

\* School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

† Ecology and Entomology Group, Soil, Plant & Ecological Sciences Division, PO Box 84, Lincoln University, Canterbury, New Zealand

A major component of human-induced global change is the deliberate or accidental translocation of species from their native ranges to alien environments<sup>1,2</sup>, where they may cause substantial environmental and economic damage<sup>3,4</sup>. Thus we need to understand why some introductions succeed while others fail. Successful introductions tend to be concentrated in certain regions<sup>2</sup>, especially islands and the temperate zone, suggesting that species-rich mainland and tropical locations are harder to invade because of greater biotic resistance<sup>1,5–9</sup>. However, this pattern could also reflect variation in the suitability of the abiotic environment at introduction locations for the species introduced<sup>13,9–11</sup>, coupled with known confounding effects of non-random selection of species and locations for introduction<sup>8,12–14</sup>. Here, we test these alternative hypotheses using a global data set of historical bird introductions, employing a statistical framework that accounts for differences among species and regions in terms of introduction success. By removing these confounding effects, we show that the pattern of avian introduction success is not consistent with the biotic resistance hypothesis. Instead,

success depends on the suitability of the abiotic environment for the exotic species at the introduction site.

Historical records provide a valuable resource for investigating reasons for introduction success or failure, but firm conclusions have proved difficult to draw from such data<sup>3,8,13</sup> for two reasons<sup>14</sup>. First, patterns of success are confounded because species were not randomly assigned to introduction locations, and because some locations will have received disproportionately more invaders, or more good or poor invaders. Second, individual introductions cannot be regarded as independent data points in a statistical analysis. Instead, introduction outcomes are likely to be correlated because the same species was frequently introduced to many locations, and because most locations were subject to several introductions<sup>15,16</sup>.

To overcome these problems, we modelled the success or failure of all known historical bird introductions using a generalized linear mixed model (GLMM), including as random effects variables that coded for the clustering of introduction events within species, higher taxa and biogeographic region of introduction (Methods). This allows us to control for differences in introduction success rate among species and regions, and to account for the non-independence of introductions by modelling the covariance among introductions of the same species to different locations, among introductions of species within higher taxa, and among introductions to the same biogeographic region.

Having controlled for the above effects, there is no relationship between introduction success and either the latitude of introduction or whether the introduction was to an island or mainland location (Table 1, fixed effects). Thus, within biogeographic regions, species-rich locations (for example, low latitude, mainland) are as easy to invade as species-poor locations (for example, high latitude, island)<sup>8,17</sup>. There was significant variation in success rate among biogeographic regions (Table 1, random effects) but the ordering of regions by their ease of invasibility was not consistent with the biotic resistance hypothesis (Table 2). Two of the most species-rich regions, the Afrotropics and Central/South America were ranked among the easiest to invade. Hence, our results find no support for

Table 1 Fixed-effect and random-effect parameters

| Fixed effect                         | Category                     | Parameter estimate | Standard error | Type III, <i>F</i>    | <i>P</i> |
|--------------------------------------|------------------------------|--------------------|----------------|-----------------------|----------|
| Biogeographic region                 | Between                      | −0.8477            | 0.1822         | 21.7                  | <0.0001  |
|                                      | Within                       | 0                  |                |                       |          |
| Latitudinal difference               |                              | −0.0366            | 0.0075         | 23.8                  | <0.0001  |
| Log <sub>10</sub> (geographic range) |                              | 0.5906             | 0.1502         | 15.5                  | <0.0001  |
| Latitude of introduction             |                              | −0.0099            | 0.0067         | 2.2                   | 0.140    |
| Island/mainland                      | Island                       | 0.0074             | 0.1743         | 0                     | 0.966    |
|                                      | Mainland                     | 0                  |                |                       |          |
| Random effect                        |                              |                    |                |                       |          |
|                                      | Estimated variance component |                    | Standard error | Likelihood ratio test | <i>P</i> |
| Among orders                         | 0.0107                       |                    | 0.0399         | 0.6                   | 0.439    |
| Among families (Order)               | 0                            |                    | —              | —                     | —        |
| Among genera (Family)                | 0.2667                       |                    | 0.2507         | 0.2                   | 0.655    |
| Among species (Genus)                | 1.5238                       |                    | 0.3201         | 226.9                 | <0.0001  |
| Among biogeographic regions          | 0.1676                       |                    | 0.1244         | 22.5                  | <0.0001  |

Parameters are estimated in a multivariate generalized linear mixed model (GLMM) with introduction success or failure as the response variable for global bird introductions. For the fixed effects, positive parameter estimates mean that larger values of the fixed effect are associated with a higher probability of introduction success, accounting for non-independence in the data owing to the clustering of introductions modelled by the random effects. For binary variables (within or between regions, island or mainland), parameter estimates are calculated relative to one of the two categories, which thus has an estimate of zero in each case. For the random effects, the likelihood ratio is a test for zero variance in the random effect with everything else in the model. Similar results are obtained for the random effects if these are modelled alone: this is effectively a nested analysis of variance (ANOVA) of introduction success across taxonomic levels<sup>15</sup>, confirming that most variation is among species within genera, and hence that the success of a species cannot be predicted from that of its relatives. The biotic resistance hypothesis predicts that introduction success should be higher on islands and lower at low latitudes<sup>1</sup>: these are both falsified. The abiotic suitability hypothesis predicts that introduction success should be higher when translocations are to similar latitudes<sup>18</sup>, within biogeographic regions<sup>2</sup>, and for species that have larger geographic ranges<sup>22</sup>: these are all supported. F, test statistic; P, probability.



an effect of biotic resistance on introduction success<sup>8,13</sup>. The traditional perception, that species-poor islands and temperate locations are easy to invade, probably reflects the greater total number of avian introductions to these locations<sup>14</sup>.

Introduction success was significantly greater when the difference between a species latitude of origin and its latitude of introduction was small, and when species were introduced to locations within their native biogeographic region (Table 1, fixed effects). Locations at similar latitude and in the same biogeographic region are more likely to share climatic and habitat features in common<sup>6,18</sup>, showing that introduction success is enhanced if species are matched with suitable environments<sup>19</sup>. All else being equal, species with larger geographical range sizes should have a wider environmental tolerance<sup>20</sup>, or use more widespread resources<sup>21</sup>, and so have a higher probability of encountering an abiotic environment at a new location that allows successful introduction<sup>22</sup>. Consistent with this prediction, geographical range size is also a significant correlate of introduction success for global bird introductions (Table 1, fixed effects). These results support the view that physiological tolerances are likely to be at least as important as biotic interactions in determining the responses of species to global climate change<sup>23,24</sup>.

Variation in introduction success unaccounted for by the fixed effects in the model can be decomposed into variation due to differences among species at different levels of the taxonomic hierarchy, differences among biogeographic regions (see above) and residual variation. This decomposition revealed significant unexplained variation among species, but not among taxonomic groups at the genus level or higher (Table 1, random effects). This shows that unmeasured species-level traits associated with introduction success must be phylogenetically labile, varying even among closely related species. By the same token, characteristics typically shared by related species, including several life history traits hypothesized to predict introduction success (body mass, generation time and population growth rate<sup>3,25–27</sup>), can explain only a small and non-significant amount of variation in global introduction success. To check this, we ran our analysis again, including family-typical values for three traits (body mass, clutch size and incubation period; see Methods) as fixed-effect predictors in the model. None of these family-level traits were significant determinants of introduction success, as predicted by the decomposition of unexplained variance among avian taxonomic levels (Table 1, random effects).

In addition to variation in specific characteristics that affect invasibility, significant differences among species in introduction success will almost certainly reflect variation in introduction effort (the number of releases or individuals released at each location)<sup>3</sup>. Although data are not available for most locations, our analyses account for variation in effort in two ways. First, species with a larger geographical range size tend to have been introduced more often and in greater numbers because they were more readily obtained for introduction<sup>14,22</sup>. Hence, the significant effect of geographical range

size on introduction success may reflect introduction effort in addition to abiotic tolerance, both of which are required to account for the significant relationship between geographical range size and success in Australian bird introductions<sup>22</sup>. Second, coding species identity as a random effect in the model controls for differences among species, including variation in introduction effort.

Determining the causes of introduction success is important for improving our ability to identify and screen out environmentally and economically damaging biological invaders<sup>3,28</sup>. Our results show that for birds the outcome of introductions is not predicted by general features of locations related to biotic resistance (such as latitude), and that the success of a species cannot be predicted from that of its relatives. Instead, the importance of phylogenetically labile species-specific factors, such as geographical range size, and event-level effects related to environmental suitability, such as the match between latitudes of origin and introduction, suggest that success depends on the particular combination of species and location. This may help to explain why general features of invaders have been hard to characterize<sup>3,13,26</sup>, but shows that an understanding of introduction success is possible nonetheless. □

## Methods

### Data

We included all known historical human-mediated introductions (excluding natural colonizations and subsequent unassisted colonizations from introduced populations) except for introductions undertaken for conservation purposes, giving a total of 1,466 introduction events (of a species to a particular location) for 389 species. We define a location as an island or a country on a continental mainland to which a species was introduced. Multiple releases to the same island or country are counted as one introduction to that location. We scored an introduction as successful if it resulted in the establishment of a persistent or probably persistent population following release, and unsuccessful otherwise (introductions described as possible successes were ignored). For each introduction event, five variables were recorded in addition to success: whether introduction was to a mainland or island location, the latitude of introduction, the difference between the latitudinal midpoint of the native range of the species and the latitude of introduction ('latitudinal difference') calculated without reference to hemisphere (that is, an introduction from 50° N to 40° S has a latitudinal difference of 10°), whether the introduction was intraregional (introduction site and native range in the same biogeographic region) or interregional (in different biogeographic regions), and the geographical range size ( $\log_{10}(\text{latitudinal range} \times \text{longitudinal range})$ ) of the species introduced. Latitude and longitude were measured in degrees. For each bird family with at least one introduced species, we calculated geometric mean values of body mass, clutch size and incubation period. All data sources are listed in the Supplementary Information.

### Analyses

We used the GLIMMIX macro in SAS<sup>29</sup> to fit a generalized linear mixed model (GLMM) specifying a binomial error distribution and logit link function, with introduction outcome (success or failure) as the response variable. GLMMs provide a framework for analysing binary response data (success or failure of introduction) in which observations are likely to be correlated owing to clustering and cannot therefore be treated as statistically independent units<sup>30</sup>. GLMMs incorporate information on the clustering to provide estimates of standard errors corrected for this non-independence, which will generally be more conservative than estimates obtained if the clustering is ignored. GLMMs also allow the variance explained at different levels of a clustering hierarchy to be decomposed, providing an upper limit to the amount of variation that could be explained by unobserved variables associated with each level of the hierarchy<sup>30</sup>. We modelled the likely non-independence of introductions of the same species by assuming a common positive correlation between introduction outcomes involving the same species, but a zero correlation between introduction outcomes involving different species (a variance components model). Clustering of introduction events within higher taxonomic levels (genera, families and orders) and biogeographic regions were similarly modelled. On the logit scale, the effects for each clustering variable are assumed to be a random sample taken from a larger population of normally distributed values, rather than being a set of fixed values, so that inferences about the importance of these 'random effects' apply to the wider population from which those observations derive. The remaining predictor variables were modelled as fixed effects.

Received 28 June; accepted 18 September 2001.

**Table 2 Parameter estimates and standard errors**

| Biogeographic region  | Parameter estimate | Standard error |
|-----------------------|--------------------|----------------|
| Afrotropics           | 0.5620             | 0.2809         |
| Palearctic            | 0.3435             | 0.2464         |
| Central/South America | 0.3098             | 0.2915         |
| Caribbean             | 0.1308             | 0.2508         |
| Malagasy              | 0.0345             | 0.2346         |
| Pacific               | -0.0210            | 0.1942         |
| Antarctic             | -0.0533            | 0.3887         |
| Australasia           | -0.2054            | 0.1949         |
| Nearctic              | -0.2516            | 0.2343         |
| Southeast Asia        | -0.2589            | 0.2677         |
| Atlantic              | -0.5905            | 0.2745         |

The biogeographic regions were included as random effects in a multivariate GLMM with introduction success or failure as the response variable for global bird introductions. Regions with more positive parameter estimates were associated with a higher probability of introduction success, accounting for the fixed effects (in Table 1) and non-independence in the data owing to the clustering of introductions by species.

1. Elton, C. *The Ecology of Invasions by Animals and Plants* (Methuen, London, 1958).
2. Vitousek, P. M., D'Antonio, C. M., Loope, L. L., Rejmánek, M. & Westbrooks, R. Introduced species: a significant component of human-caused global change. *NZ J. Ecol.* **21**, 1–16 (1997).
3. Williamson, M. *Biological Invasions* (Chapman & Hall, London, 1996).
4. Mack, R. N. *et al.* Biotic invasions: causes, epidemiology, global consequences, and control. *Ecol. Appl.* **10**, 689–710 (2000).
5. Pimm, S. L. in *Biological Invasions, a Global Perspective* (eds Drake, J. *et al.*) 351–367 (Wiley, Chichester, UK, 1989).

6. Simberloff, D. in *Biological Invasions, a Global Perspective* (eds Drake, J. et al.) 61–75 (Wiley, Chichester, UK, 1989).
7. Case, T. J. Invasion resistance, species build up, and community collapse in metapopulation models with inter-species competition. *Biol. J. Linn. Soc.* **42**, 239–266 (1991).
8. Rejmánek, M. in *Biodiversity and Ecosystem Processes in Tropical Forests* (eds Orians, G. H., Dirzo, R. & Cushman, J. H.) 153–172 (Springer, Berlin, 1996).
9. Holdgate, M. W. Summary and conclusions: characteristics and consequences of biological invasions. *Phil. Trans. R. Soc. Lond. B* **314**, 733–742 (1986).
10. Newsome, A. E. & Noble, I. R. in *Ecology of Biological Invasions* (eds Groves, R. H. & Burdon, J. J.) 1–20 (Cambridge Univ. Press, Cambridge, 1986).
11. Brown, J. H. in *Biological Invasions, a Global Perspective* (eds Drake, J. A. et al.) 85–109 (Wiley, Chichester, UK, 1989).
12. Crawley, M. J. The population biology of invaders. *Phil. Trans. R. Soc. Lond. B* **314**, 711–731 (1986).
13. Lonsdale, W. M. Global patterns of plant invasions and the concept of invasibility. *Ecology* **80**, 1522–1536 (1999).
14. Blackburn, T. M. & Duncan, R. P. Establishment patterns of exotic birds are constrained by non-random patterns in introduction. *J. Biogeogr.* (in the press).
15. Harvey, P. H. & Pagel, M. D. *The Comparative Method in Evolutionary Biology* (Oxford Univ. Press, Oxford, 1991).
16. Case, T. J. Global patterns in the establishment and distribution of exotic birds. *Biol. Conserv.* **78**, 69–96 (1996).
17. Sol, D. Are islands more susceptible to be invaded than continents? Birds say no. *Ecography* **23**, 687–692 (2000).
18. Crawley, M. J., Harvey, P. H. & Purvis, A. in *Plant Life Histories: Ecology, Phylogeny and Evolution* (eds Silvertown, J., Franco, M. & Harper, J. L.) 36–52 (Cambridge Univ. Press, Cambridge, 1997).
19. Williamson, M. & Brown, K. C. The analysis and modelling of British invasions. *Phil. Trans. R. Soc. Lond. B* **314**, 505–522 (1986).
20. Brown, J. H. On the relationship between abundance and distribution of species. *Am. Nat.* **124**, 255–279 (1984).
21. Gaston, K. J. *Rarity* 117–124 (Chapman & Hall, London, 1994).
22. Duncan, R. P., Bomford, M., Forsyth, D. M. & Conibear, L. High predictability in introduction outcomes and the geographical range size of introduced Australian birds: a role for climate. *J. Anim. Ecol.* **70**, 621–632 (2001).
23. Davis, A. J., Jenkinson, L. S., Lawton, J. H., Shorrocks, B. & Wood, S. Making mistakes when predicting shifts in species range in response to global warming. *Nature* **391**, 783–786 (1998).
24. Hodkinson, I. D. Species response to global environmental change or why ecophysiological models are important: a reply to Davis et al. *J. Anim. Ecol.* **68**, 1259–1262 (1999).
25. Lawton, J. H. & Brown, K. C. The population and community ecology of invading insects. *Phil. Trans. R. Soc. Lond. B* **314**, 606–617 (1986).
26. Crawley, M. J. in *Colonization, Succession and Stability: 26th Symposium of the British Ecological Society* (eds Gray, A. J., Crawley, M. J. & Edwards, P. J.) 429–453 (Blackwell Scientific, Oxford, 1987).
27. Ehrlich, P. R. in *Biological Invasions, a Global Perspective* (eds Drake, J. A. et al.) 315–328 (Wiley, Chichester, UK, 1989).
28. Smith, C. S., Lonsdale, W. M. & Fortune, J. When to ignore advice: invasion predictions and decision theory. *Biol. Invas.* **1**, 89–96 (1999).
29. Littell, R. C., Milliken, G. A., Stroup, W. W. & Wolfinger, R. D. *SAS System for Mixed Models* (SAS Institute, Cary, North Carolina, 1996).
30. Goldstein, H. *Multilevel Statistical Models* (Edward Arnold, London, 1995).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

## Acknowledgements

We thank J. Baker, J. Bale, J. Buonaccorsi, C. Frampton and R. Kinnersley for comments on the manuscript, and the University of Birmingham for funding assistance.

Correspondence and requests for materials should be addressed to T.M.B. (e-mail: t.blackburn@bham.ac.uk).

# Distance determined by the angular declination below the horizon

Teng Leng Ooi\*, Bing Wu† & Zijiang J. He†

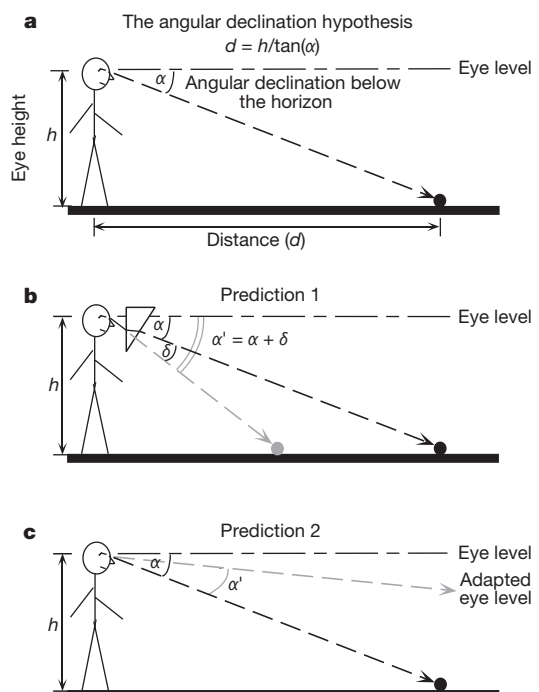
\* Department of Biomedical Sciences, Southern College of Optometry, Memphis, Tennessee 38104, USA

† Department of Psychological and Brain Sciences, University of Louisville, Louisville, Kentucky 40292, USA

A biological system is often more efficient when it takes advantage of the regularities in its environment<sup>1,2</sup>. Like other terrestrial creatures, our spatial sense relies on the regularities associated with the ground surface<sup>2–6</sup>. A simple, but important, ecological fact is that the field of view of the ground surface extends upwards

from near (feet) to infinity (horizon)<sup>2</sup>. It forms the basis of a trigonometric relationship wherein the further an object on the ground is, the higher in the field of view it looks, with an object at infinity being seen at the horizon. Here, we provide support for the hypothesis that the visual system uses the angular declination below the horizon for distance judgement. Using a visually directed action task<sup>7–10</sup>, we found that when the angular declination was increased by binocularly viewing through base-up prisms, the observer underestimated distance. After adapting to the same prisms, however, the observer overestimated distance on prism removal. Most significantly, we show that the distance overestimation as an after-effect of prism adaptation was due to a lowered perceived eye level, which reduced the object's angular declination below the horizon.

Figure 1a illustrates the relationship between the angular declination below the horizon ( $\alpha$ ) and the absolute distance ( $d$ ) of an object on the ground from the observer. Assuming that the observer's eye height ( $h$ ) is known<sup>11</sup>, the object distance can be determined by obtaining the angular declination below the horizon:  $d = h/\tan(\alpha)$  (refs 2–4). To test this 'angular declination hypothesis', that the visual system can access the information regarding the angular declination below the horizon for distance perception, consider the consequence of viewing through a pair of base-up prisms that deviate light by  $\delta$  degrees (Fig. 1b). Predictably, the angular declination below the horizon will increase to  $\alpha + \delta$ , and, accordingly, the perceived distance,  $h/\tan(\alpha + \delta)$ , will decrease. Next, suppose that the observer continually views the visual environment through the base-up prisms, and eventually removes the prisms to reveal the after-effect of prism adaptation<sup>12–14</sup>. If we assume that prism adaptation induces a recalibration of the eye level downward, the angular declination below the horizon will also reduce. Thus, we



**Figure 1** The angular declination hypothesis. **a**, The hypothesis describes how the visual system can compute distance ( $d$ ) from the eye height ( $h$ ) and the angular declination below the horizon ( $\alpha$ ), with the trigonometric relationship  $d = h/\tan(\alpha)$ . **b**, Prediction 1. A base-up prism increases the object's angular declination from  $\alpha$  to  $\alpha + \delta$ , reducing computed distance. **c**, Prediction 2. The after-effect of base-up prism adaptation is a downward shift of the eye level. Because the eye level serves as a reference for computing the angular declination below the horizon, and the object's angular declination ( $\alpha$ ) is reduced to ( $\alpha'$ ), distance is overestimated.

predict that the perceived distance will increase (Fig. 1c).

We tested both predictions using a blindfolded-walking paradigm<sup>7–10</sup>, where an observer binocularly previews a target and then walks to the target's location in blindfold traversing the remembered target distance (a visually directed task). In our experiments, which were conducted in a well-lit visual environment, we first measured seven naive observers' baseline performances. We found that the observers performed the task quite accurately (Fig. 2a). Then, we measured the observers as they viewed through a pair of 10 Prism Diopters (PD) (5.73 degrees) base-up prisms. This time, they significantly underestimated distance (Fig. 2a) compared with the baseline condition (two-way analysis of variance (ANOVA) with repeated measures,  $F = 41.053$ , degrees of freedom d.f. = 1,6,  $P < 0.001$ ). Thus, our finding confirms the first prediction of the angular declination hypothesis (Fig. 1b).

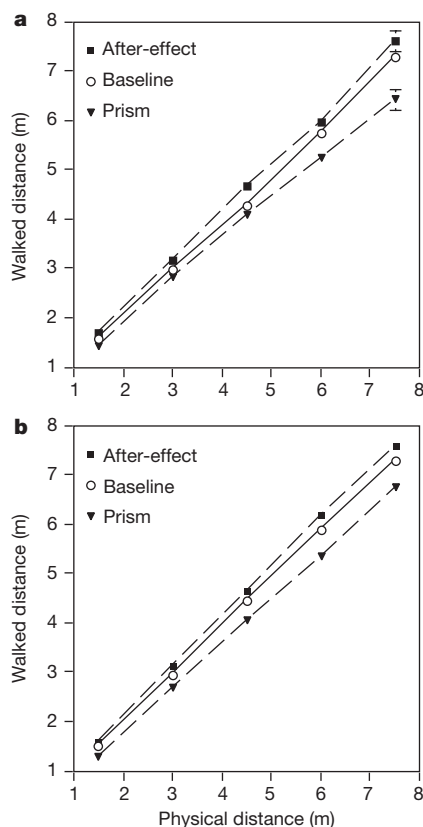
To test the second prediction (Fig. 1c), we induced prism adaptation in the same observers by having them perform purposeful walking while wearing the pair of 10 PD base-up prisms for 20 min. Thereafter, we removed the prisms and measured the observers' distance judgements (after-effect condition) using the blindfolded-walking paradigm. Their averaged walked distances are shown in Fig. 2a, which reveals distance overestimations compared with the baseline (two-way ANOVA with repeated measures,  $F = 24.16$ , d.f. = 1,6,  $P < 0.003$ ), confirming our second prediction.

The second prediction was reconfirmed by a different prism-adaptation method. We adapted the same observers by having them stand still and throw beanbags to two targets on the floor (1.8 and 3.6 m) while wearing the pair of 10 PD base-up prisms for 20 min. We then measured the prism after-effect, and found the observers

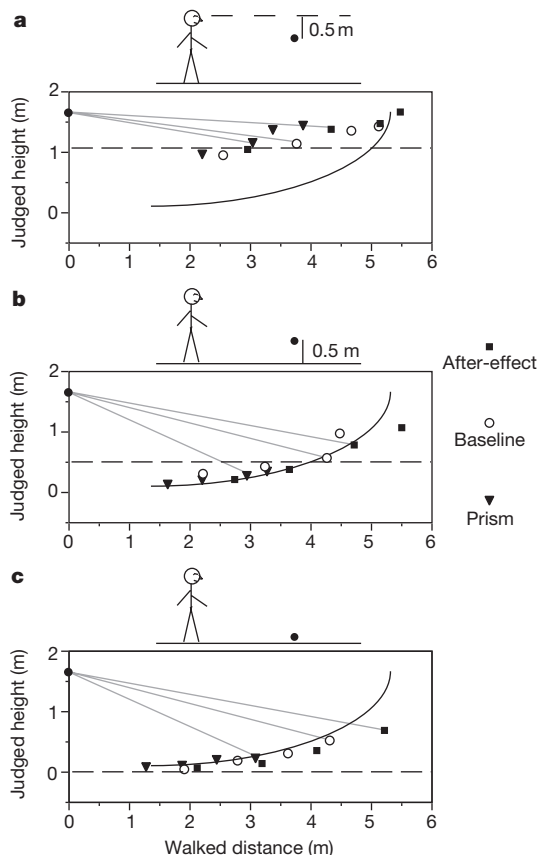
overestimated target distances (two-way ANOVA with repeated measures,  $F = 11.65$ , d.f. = 1,6,  $P < 0.015$ ). Their results are shown in Fig. 2b, which also plots their performances in the baseline and prism conditions that were tested before the throwing-adaptation experiment. Clearly, these results are similar to those in the walking-adaptation experiment (Fig. 2a). Importantly, because the observers did not walk during the throwing-adaptation phase, we can rule out the possibility that the after-effect found with the walking-adaptation experiment was solely due to an adaptation within the walking-locomotion system (a blindfolded-walking paradigm was used to measure the perceived distance)<sup>15</sup>.

Our findings have confirmed the two predictions of the angular declination hypothesis. This hypothesis, however, hinges on the implicit assumption that the visual system uses the eye level as a reference for computing the angular declination of the object (Fig. 1a). But while this assumption is critical, it also lacks direct empirical support. Thus, the remainder of our study was dedicated to proving that a recalibration of the eye level serves as the mechanism for deriving distance judgements, from the angular declination below the horizon. Below, we measured the impact of base-up prism adaptation on judgements of eye level and target location. Both sets of experiments were conducted in the dark to increase the reliability of the eye level measurements<sup>16</sup>.

We first measured the observer's eye level by having the observer instruct the experimenter to locate a red light at the observer's eye level from a viewing distance of 2.4 m. Six new naive observers and



**Figure 2** The effect of viewing through a pair of 10 PD base-up prisms and the after-effect of adaptation to base-up prisms on judged distance in the light ( $n = 7$ ). Relative to the no-prism baseline condition, observers underestimated distance when viewing through prisms, and overestimated distance after prism adaptation when measured with prisms removed. Prism adaptation was induced by (a) purposeful walking and (b) throwing beanbags.

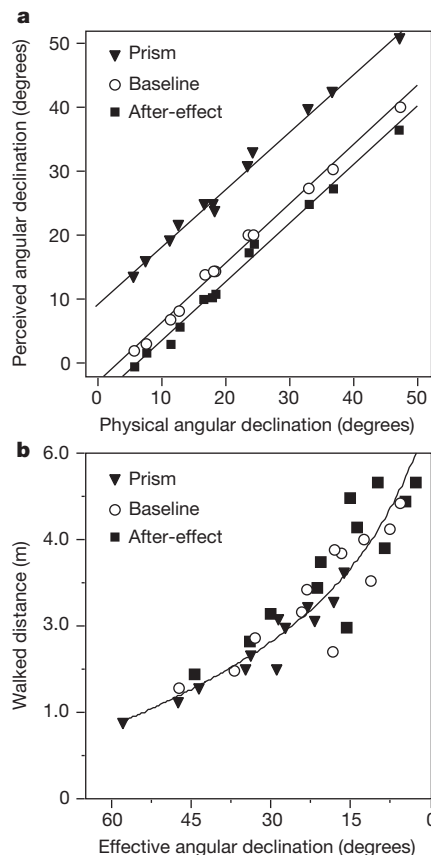


**Figure 3** Perceived locations of a dim target in the dark for the baseline, prism and after-effect conditions. The target heights (indicated by the dashed lines) were 0.5 m below eye level (a); 0.5 m above the floor (b); or on the floor (c). Relative to baseline, judged target location was further and higher in the after-effect condition, and nearer and lower in the prism condition ( $n = 7$ ). This general trend is emphasized in each graph by the three grey lines, which connect the perceived target locations to the average eye positions on the y axes (for the same test target in the three viewing conditions). Doing so allows us to calculate the perceived angular declination for each judged target location (Fig. 4).



one author participated in all three conditions: baseline, prism and after-effect. To induce prism adaptation, the observers wore a pair of 20 PD (11.5 degrees) base-up prisms and walked about in the light for 20 min. Consistent with previous reports<sup>17</sup>, the judged eye level was slightly below the physical eye level ( $-2.19 \pm 0.57$  degrees, mean  $\pm$  s.d.) in the baseline condition. For the prism condition, the eye level was higher than baseline ( $8.20 \pm 1.89$  degrees), whereas it was lower than baseline in the after-effect condition ( $-5.00 \pm 0.92$  degrees). Clearly, the results indicate that the eye level can be altered by prisms.

In another experimental session, we measured the same seven observers' judgements of target locations in the dark. During a trial, the observer previewed a red light target that was randomly placed at 1 of 12 predetermined locations (4 distances  $\times$  3 heights). The observer walked the remembered target distance and gestured the target's height. Figure 3 shows the average results in three separate graphs for the three target heights tested. The graphs relate the judged target height ( $y$  axis) to the judged target distance ( $x$  axis), respectively, for the baseline, prism and after-effect conditions. Also indicated on the  $y$  axes are the average physical eye heights, which we connected by lines to a representative datum from each of the three viewing conditions. This allows us to illustrate the changes in perceived angular declination across viewing conditions. Overall,



**Figure 4** Veridical judgement of angular declination and its monotonic relationship to distance. **a**, The perceived angular declination derived from Fig. 3 plotted as a function of physical angular declination below the horizon, along with the regression lines for each viewing condition. For the baseline condition, the slope of the regression line is 0.922. The slopes for the prism and after-effect conditions are 0.902 and 0.910, respectively. They are not significantly different from the baseline condition ( $F = 0.411$ , d.f. = 1, 11,  $P = 0.54$ ;  $F = 0.275$ , d.f. = 1, 11,  $P = 0.61$ ). **b**, A monotonic increase in walked distance with effective angular declination. The hyperbolic curve  $d = h/[a \tan(\alpha) + b]$  was fitted to the data by the least squares method (where  $d$  is distance,  $\alpha$  is effective angular declination and  $h$  is average eye height (1.62 m); constant  $a = 0.6194$  and constant  $b = 0.2426$ ).

compared to baseline, angular declination is increased in the prism condition, resulting in significant decreases in perceived target distances ( $t = 10.17$ , d.f. = 11,  $P < 10^{-6}$ ) and heights ( $t = 4.66$ , d.f. = 11,  $P < 0.001$ ). Meanwhile, angular declination is decreased in the after-effect condition, resulting in significant increases in perceived target distances ( $t = 7.99$ , d.f. = 11,  $P < 10^{-5}$ ) and heights ( $t = -2.63$ , d.f. = 11,  $P < 0.025$ ).

The data in Fig. 3b seem to be distributed along a trend, about the curve, which we fitted by eye. We also drew the same curve in Fig. 3a and c. As can be seen, the fit is good for Fig. 3c. Thus, Fig. 3b and c suggest that in the dark, the visual system tends to treat targets on or near the floor as if they are located on a distinct curved surface<sup>2,18</sup>. For Fig. 3a, only data points representing further target distances fit the curve, whereas those representing nearer target distances are distributed above the curve. This is probably because of their relatively short distances from the eyes, causing their distance judgement to be affected by other near depth cues such as ocular motor, binocular disparity and motion parallax cues.

We next verified the possibility that prism adaptation affects the judgements of both distance and height, possibly stemming from a recalibration of the eye level. We calculated the perceived angular declination for the data in Fig. 3, and plotted them in Fig. 4a as a function of the physical angular declination of the targets used. A linear relationship is revealed between the perceived and physical angular declination below the horizon for all three conditions. Furthermore, the three regression lines are parallel, suggesting a constant change in angular declination. We also calculated the average difference in perceived angular declination between the after-effect and baseline conditions:  $-2.90 \pm 0.44$  degrees, which is very close to the average difference in eye level ( $-2.82 \pm 0.74$  degrees) between the after-effect and baseline conditions from the eye level judgement experiments ( $t = 0.796$ , d.f. = 6,  $P = 0.46$ ). Thus, it indicates that a change in the eye level very probably affects the perceived target locations (Fig. 3). This provides compelling support for the implicit assumption of the angular declination hypothesis, that the eye level serves as a reference for computing the angular declination below the horizon.

We then performed the same analyses for the prism condition relative to baseline, wherein we found the average difference in angular declination to be  $11.63 \pm 1.37$  degrees, and the average difference in eye level to be  $10.38 \pm 1.43$  degrees. These also do not differ from each other ( $t = 0.103$ , d.f. = 6,  $P = 0.92$ ), indicating a common basis. The differences in perceived angular declination and in eye level due to the prisms are close to 11.5 degrees, which is the physical angular deviation caused by the prisms. Such concordance legitimizes the methods we employed to obtain the measurements of eye level and angular declination.

Knowing the changes in eye levels allowed us directly to relate the walked distance to the effective angular declination for each condition. (Effective angular declination for the prism condition = physical angular declination + 10.38 degrees; effective angular declination for after-effect condition = physical angular declination - 2.82 degrees). This is plotted in Fig. 4b, which shows a monotonic function, confirming that angular declination is a depth cue<sup>2</sup> for distance judgement.

That eye level is not fixed but subservient to the visual environment after a short period of prism adaptation suggests the existence of a plastic mechanism that calibrates the intrinsic body information to the visual environment<sup>12-14</sup>. Indeed, such a learning mechanism is critical for ensuring that we correctly perceive the lawful relationship between the physical properties of the visual environment and our visuomotor system<sup>12-15</sup>.

Our study provides direct support for the angular declination hypothesis, that the visual system can access the angular declination below the horizon to determine absolute distance. This finding is consistent with several empirical studies<sup>19-21</sup>, such as that by Philbeck and Loomis<sup>21</sup> in their reduced cue condition. Our

study reinforces this notion by relating the changes in the angular declination below the horizon to perceived eye level, and demonstrating that the eye level serves as a reference for the visual system to compute the angular declination below the horizon. □

## Methods

### Observers

Thirteen naive observers with informed consent and one author with self-reported normal vision participated in the various experiments.

### Blindfolded-walking paradigm in the light

The observer stood in a hallway 25 m long (starting about 10 m to one end of the hallway) and previewed a rectangular cardboard target (2.12 degrees) on the floor at one of five distances (1.5, 3.0, 4.5, 6.0 or 7.5 m). The observer then wore a blindfold and walked the remembered distance in the direction of the target (which had been removed). When the judged distance was reached, the experimenter measured it, and then led the observer in blindfold to the starting point to begin a new trial (no feedback was provided). Each target distance was measured twice (counterbalanced).

### Blindfolded-walking paradigm in the dark

The observer stood in a dark room and previewed an internally illuminated red table-tennis ball (target, 0.16 cd m<sup>-2</sup>) at one of four distances (1.5, 2.5, 3.75 or 5.0 m) and three elevations (on the floor, 0.5 m above the floor, or 0.5 m below the eye). The target was then removed for the observer to begin walking in the dark, traversing the remembered target distance. On reaching the destination, the observer gestured the remembered target elevation with his/her left hand. No feedback was given to the observer. Each target location was randomly selected and measured twice. All targets viewed from the same distance had the same physical size, which subtended 0.23 degrees at the eye level.

### Prism-adaptation method

The observer wore a pair of base-up prism goggle (10 PD (5.73 degrees) or 20 PD (11.5 degrees) from Bernell/USO) for 20 min, while actively performing one of two tasks: walking about, with the specific instruction to navigate complex obstacle courses in the natural visual environment; or standing still and throwing beanbags to a target on the floor 1.8 or 3.6 m using the right hand. To maintain the after-effect of prism adaptation during our lengthy experiment when measuring target locations in the dark, the observer wore the prisms in a lighted room between trials.

### Judging eye-level task

In a dark room, the observer stood still with his/her head held by a head and chin rest. A red light target (supported by a wall) that subtended 0.23 degrees was moved by the experimenter from a viewing distance of 2.4 m from the observer, who would instruct the experimenter to stop moving the light when it was perceived to be at the observer's eye level. The procedure was repeated five times for each condition tested.

Received 20 July; accepted 20 September 2001.

1. Attneave, F. Informational aspects of visual perception. *Psychol. Rev.* **61**, 183–193 (1954).
2. Gibson, J. J. *The Perception of the Visual World* (Houghton Mifflin, Boston, 1950).
3. Sedgwick, H. A. in *Human and Machine Vision* (eds Rosenthal, A. & Beck, J.) 425–458 (Academic, New York, 1983).
4. Sedgwick, H. A. in *Handbook of Perception and Human Performance* (eds Boff, K. R., Kaufman, L. & Thomas, J. P.) 21.1–21.57 (Wiley, New York, 1986).
5. Sinai, M. J., Ooi, T. L. & He, Z. J. Terrain influences the accurate judgement of distance. *Nature* **395**, 497–500 (1998).
6. Meng, J. C. & Sedgwick, H. A. Distance perception mediated through nested contact relations among surface. *Percept. Psychophys.* **63**, 1–15 (2001).
7. Thomson, J. A. Is continuous visual monitoring necessary in visually guided locomotion? *J. Exp. Psychol. Hum. Percept. Perform.* **9**, 427–443 (1983).
8. Rieser, J. J., Ashmead, D., Talor, C. & Youngquist, G. Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception* **19**, 675–689 (1990).
9. Loomis, J., DaSilva, J., Fujita, N. & Fukusima, S. Visual space perception and visually directed action. *J. Exp. Psychol.* **18**, 906–921 (1992).
10. Loomis, J., DaSilva, J., Philbeck, J. W. & Fukusima, S. Visual perception of location and distance. *Curr. Dir. Psychol. Sci.* **5**, 72–77 (1996).
11. Warren, W. H. & Whang, S. Visual guidance of walking through apertures: Body-scaled information for affordances. *J. Exp. Psychol. Hum. Percept. Perform.* **13**, 371–383 (1987).
12. Held, R. & Freedman, S. Plasticity in human sensorimotor control. *Science* **142**, 455–462 (1963).
13. Hay, J. C. & Pick, H. L. Jr Visual and proprioceptive adaptation to optical displacement of the visual stimulus. *J. Exp. Psychol.* **71**, 150–158 (1966).
14. Sugita, Y. Global plasticity in adult visual cortex following reversal of visual input. *Nature* **380**, 523–526 (1996).
15. Rieser, J. J., Pick, H. L. Jr, Ashmead, D. H., & Garing, A. E. Calibration of human locomotion and models of perceptual-motor organization. *J. Exp. Psychol. Hum. Percept. Perform.* **21**, 480–497 (1995).
16. Martin, L. et al. Oculoparalytic illusion: Visual-field dependent spatial mislocalization by humans partially paralyzed with curare. *Science* **216**, 198–201 (1982).
17. Stoper, A. & Cohen, M. Judgments of eye level in light and in darkness. *Percept. Psychophys.* **40**, 311–316 (1986).
18. Gogel, W. C. & Tietz, J. D. Absolute motion parallax and the specific distance tendency. *Percept. Psychophys.* **13**, 284–292 (1973).
19. Epstein, W. Perceived depth as a function of relative height under three background conditions. *J. Exp. Psychol.* **72**, 335–338 (1966).
20. Wallach, H. & O'Leary, A. Slope of regard as a distance cue. *Percept. Psychophys.* **31**, 145–148 (1982).
21. Philbeck, J. W. & Loomis, J. M. Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *J. Exp. Psychol. Hum. Percept. Perform.* **23**, 72–85 (1997).

## Acknowledgements

We thank K. A. May and P. J. Gunther for their assistance in collecting the data in the first experiment. This research was supported in part by grants from the Knights Templar Eye Foundation and the Southern College of Optometry (to T.L.O.) and a Competitive Enhancement Grant from the University of Louisville (to Z.J.H.).

Correspondence and requests for materials should be addressed to T.L.O. (e-mail: tloui@sco.edu).

# Linear processing of spatial cues in primary auditory cortex

Jan W. H. Schnupp, Thomas D. Mrsic-Flogel & Andrew J. King

University Laboratory of Physiology, University of Oxford, Parks Road, Oxford OX1 3PT, UK

To determine the direction of a sound source in space, animals must process a variety of auditory spatial cues, including interaural level and time differences, as well as changes in the sound spectrum caused by the direction-dependent filtering of sound by the outer ear<sup>1</sup>. Behavioural deficits observed when primary auditory cortex (A1) is damaged have led to the widespread view that A1 may have an essential role in this complex computational task<sup>2–5</sup>. Here we show, however, that the spatial selectivity exhibited by the large majority of A1 neurons is well predicted by a simple linear model, which assumes that neurons additively integrate sound levels in each frequency band and ear. The success of this linear model is surprising, given that computing sound source direction is a necessarily nonlinear operation<sup>6–9</sup>. However, because linear operations preserve information, our results are consistent with the hypothesis that A1 may also form a gateway to higher, more specialized cortical areas<sup>10,11</sup>.

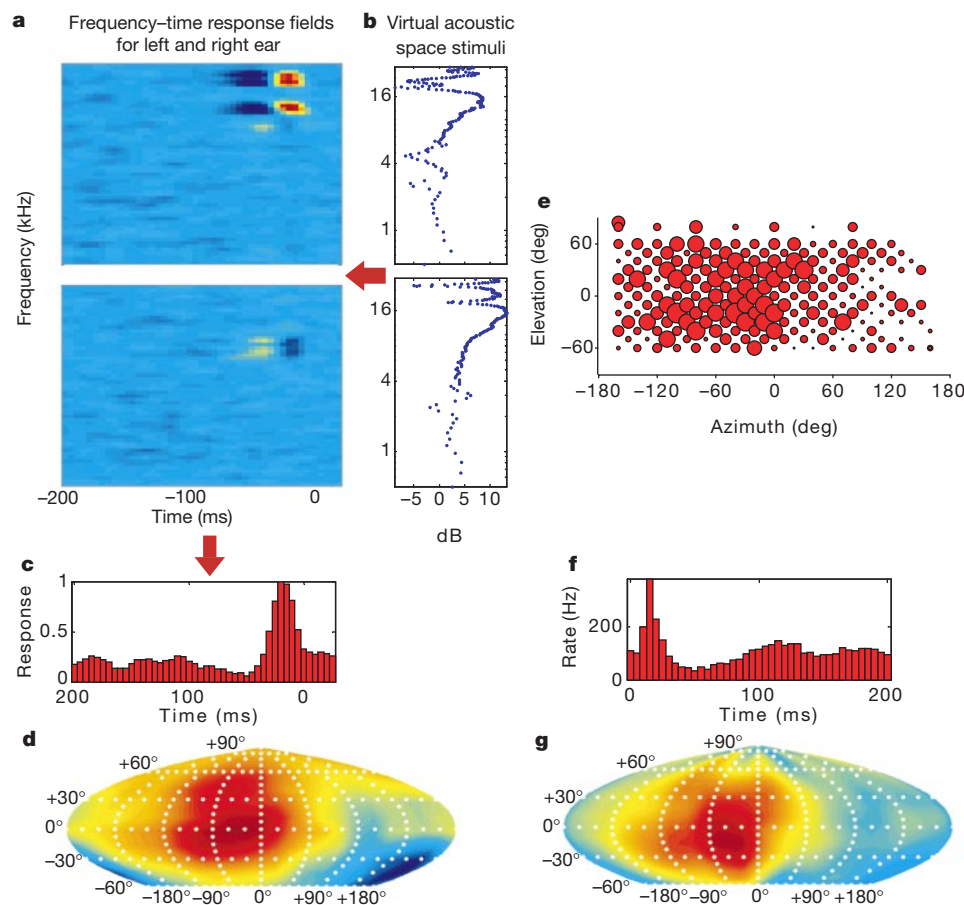
The ability of many species to pinpoint the direction of a sound source both in azimuth and elevation with an error that may be as little as a few degrees is remarkable, considering the complex and ambiguous nature of the acoustic information available to solve this task. Animals must combine spatial information provided by different acoustic cues, including time-of-arrival and level differences in the signal received at each ear, as well as changes in the spectrum of the signal generated by direction-dependent filtering of the sound by the head and external ears. Individually, these cues suffer from inherent ambiguities. For example, within a given narrow frequency band, a number of source directions can generate identical interaural time and level difference values<sup>12</sup>. Similarly, the filtering by the external ears convolves, and thus confounds, spectral localization cues with the source spectra<sup>13,14</sup>. Consequently, the central auditory system must extract, process and combine information over many frequency channels and from both ears to achieve the highly accurate localization performance exhibited by many species. The fact that damage to A1 in mammals produces marked, and specific, deficits in auditory localization performance<sup>2–5</sup>, may mean that A1 has a critical role in the pathways underlying spatial hearing and may be responsible for performing many of the computations that underlie the perception of sound source location. Indeed, the most clearly documented behavioural impairment observed following a unilateral lesion of the auditory cortical areas, including A1, is an inability to localize sounds on the side opposite the lesion<sup>2–5,15</sup>.

However, more recent studies, carried out in both awake<sup>16</sup> and anaesthetized<sup>17</sup> animals, have indicated that at least a proportion of neurons in A1 might be adequately described as linear filter elements that perform simple additive operations within their frequency–time response fields (FTRFs). The response properties of these neurons may therefore be analogous to those of simple cells in the primary visual cortex<sup>18</sup>. The FTRFs of A1 neurons have been measured using reverse correlation to stimuli presented either in the free field<sup>16</sup> or through an earphone to the contralateral ear (the ear on the opposite side of the cerebral hemisphere from which recordings were made) only<sup>17</sup>. Consequently, these studies were not able to estimate the relative contributions of inputs to each ear, or to derive predictions for responses to binaural stimulation (stimulation of both ears). Therefore, to measure binaural FTRFs of neurons in A1 of the anaesthetized ferret, we have extended this reverse correlation technique<sup>16</sup>. We concentrated on the dorsal 80% of A1 where frequencies over 4 kHz are represented<sup>19</sup>, because we are primarily interested in the role of spectral cues to sound localization. In ferrets, directional filtering by the structures of the external ear has little effect on frequencies below about 5 kHz and produces very marked spectral patterns only for frequencies of approximately 10 kHz and above<sup>20</sup>.

Using the binaural FTRFs and measurements of the acoustic localization cues generated by the animal's head and external ears we generated predictions for the shape of each neuron's spatial

response field (SRF). These were then compared to detailed maps of the actual SRFs of the same neurons, which were recorded using individualized virtual acoustic space (VAS) stimuli. Figure 1 illustrates how the observed and predicted SRF maps were produced. Sixty-nine neurons (single units) from four animals were characterized in this manner. The data shown in Fig. 1d and Fig. 2 indicate that the linear model was often highly successful at predicting the shape of the SRF (although there were some notable exceptions, see Fig. 2f). The match between the observed and predicted SRF was quantified by calculating the correlation coefficient for observed versus predicted responses over all those stimulus directions that evoked a measurable (supra-threshold) response in the observed SRF. Figure 3 shows the distribution of correlation coefficients obtained.

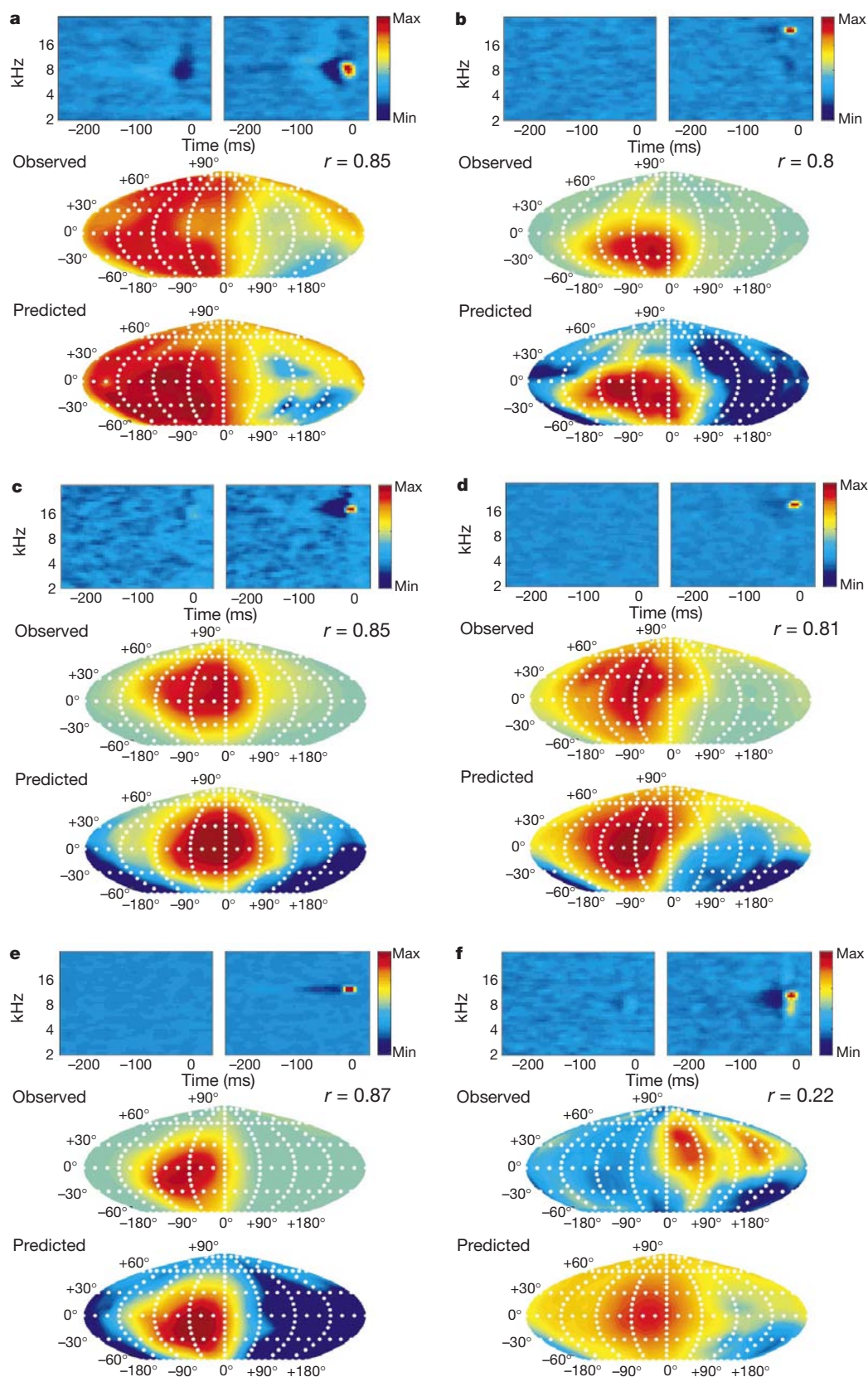
In a few cases (Fig. 2f) the model clearly failed to predict the observed SRF, indicating that some high-frequency neurons behaved in a highly nonlinear manner. The predictions also failed for the very few units we recorded which had best frequencies  $f_B$  well below 4 kHz. This is probably because our binaural FTRFs cannot reveal ITD sensitivity, which can occur in A1 neurons responsive to these relatively low frequencies<sup>21</sup>. However, 51/69 (74%) of the units with  $f_B > 4$  kHz gave  $r$  values (correlation coefficients) above 0.6 and almost half (33/69) gave  $r > 0.7$ . These values, as well as the examples in Fig. 2a–e, illustrate that the linear model produced, in most cases, a close approximation of the spatial response profile for



**Figure 1** Predicting spatial responses from frequency-filtering characteristics. **a**, Frequency–time response fields (FTRFs) of A1 neurons measured by reverse correlation (see Methods). Convolution (matrix multiplying) the FTRFs with the energy spectrum vectors of the virtual acoustic space (VAS) stimuli for each ear (**b**) generates a (time-reversed) prediction for the response post stimulus time histogram (PSTH). Summing the predicted PSTHs over all VAS stimuli we obtain a ‘pooled prediction’ (in arbitrary units) (**c**) which resembles the observed pooled PSTH (**f**) obtained from the actual

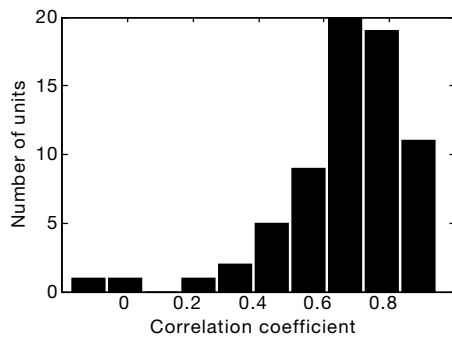
responses to VAS stimuli. **d**, Predicted spatial response field (SRF) generated by measuring the size of the predicted response peak for each VAS position in turn. **e**, ‘Raw’ SRF, measured with VAS stimuli; the diameter of the red circles is proportional to the mean response. **g**, Observed SRF map after interpolation and smoothing. Stimulus levels were as follows. VAS (dB above unit threshold): 25; FTRF-weighted VAS (dB SPL): 71; Single tonal reverberant component (dB SPL): 60 (see Methods).





**Figure 2** Examples of binaural FTRFs, shown alongside observed and predicted SRFs for six units. Five units generated good predictions (**a–e**) and for one unit the prediction failed (**f**). As none of the units illustrated here exhibited significant patterns in their binaural FTRFs at frequencies below 2 kHz, only the top four octaves of the measured binaural

FTRFs are shown. Correlation ( $r$  values) between observed and predicted responses are also given. Stimulus levels for **a** to **f**, respectively were as follows. VAS (dB above unit threshold): 20, 40, 20, 25, 15 and 25; FTRF-weighted VAS (dB SPL): 31, 70, 25, 27, 25 and 38; single tonal revcor component (dB SPL): 50, 46, 50, 40, 35 and 50.



**Figure 3** Distribution of correlation coefficients for predicted versus observed spatial response field maps.

high-frequency neurons in A1.

The use of VAS stimuli that mimic the acoustic properties of the head and outer ears of other individuals, rather than the animal's own, can lead to marked changes in the shapes of the observed SRF<sup>22</sup>. When we were able to record the binaural FTRF as well as own-ear and foreign-ear SRFs for the same unit, we invariably found that the changes produced by switching to foreign-ear VAS stimuli were also well predicted by the linear model. Figure 4 shows one example for which the differences between the own-ear and foreign-ear SRFs were particularly conspicuous. The successful prediction of the SRFs measured using both sets of VAS stimuli provides a further illustration of the fact that most A1 neurons are well approximated by a binaural linear filter model.

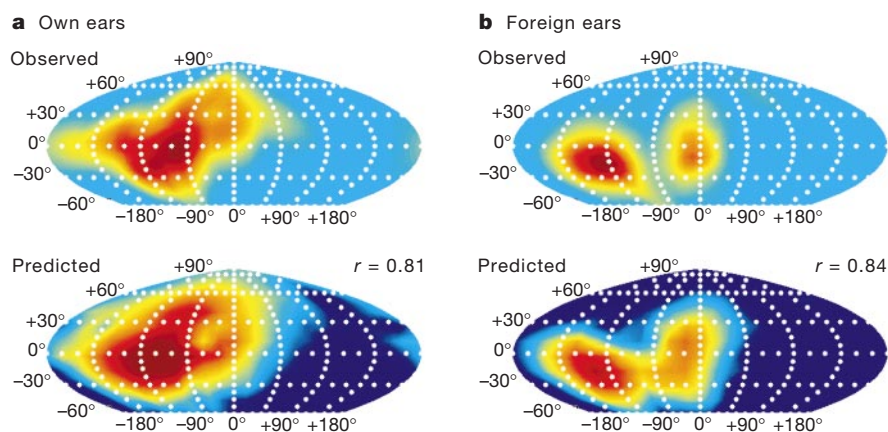
The behaviour of all but the simplest physical systems is at best only approximately linear over a limited range of input intensities. Nevertheless, linear models can be of great value if they produce good approximations for a sufficiently large range of inputs. In this study, no attempt was made to match the sound level of the VAS stimuli used to map SRFs to that of the random chord sequences used to estimate binaural FTRFs. VAS stimulus levels within the excitatory bands revealed by the binaural FTRFs were estimated after the measurements and found to be distributed randomly, with a  $\pm 40$  dB range, relative to the level of the random chord stimuli (see Methods). The fact that the linear model generated useful predictions despite these variations in level gives some confidence that the model holds over a range of sound levels. Also, in a small number of cases, we were able to repeat the SRF mapping at two or three different sound levels. In keeping with previous studies of A1 (refs 23–25), we found that the size and shape of the observed SRFs varied with the sound level used, the  $f_B$  of the units and the nature of

their binaural response properties. In general, units with high  $f_B$  values and with excitatory inputs from the contralateral ear and inhibitory inputs from the ipsilateral ear (the ear on the same side as the cerebral hemisphere from which recordings were taken) showed the sharpest spatial tuning. At sound levels very close to the neurons' threshold, the observed SRF typically covered just the tip of the predicted SRF shape, whereas, with increasing level, progressively more of the predicted SRF exceeded the threshold and became visible in the observed SRF.

Marked nonlinearities have been described at practically all levels of the auditory pathway, beginning with the mechanical response of the basilar membrane<sup>26</sup>. As in other parts of the central auditory pathway, responses of A1 neurons are known to exhibit a number of well documented nonlinearities, including non-monotonic rate-level functions and binaural facilitation or occlusion (reviewed in ref. 21). Against this background, the good performance of our linear, binaural filter model may seem surprising, particularly as linear operations have little computational power. Nonlinear operations, in contrast, help neurons develop highly selective and specific tuning properties, and underlie the sensitivity of auditory neurons to acoustic localization cues<sup>6–8</sup> and to combinations of those cues<sup>9</sup>.

If one of the main functions of A1 is indeed to compute the location of a sound source<sup>2–5</sup>, we might expect neurons in this cortical field to make extensive use of the power inherent in nonlinear operations. On the other hand the amount of information a highly selective neuron transmits on average about the sensory environment will be reduced, because it conveys no information at all about the many aspects of the environment to which its selective properties have rendered it insensitive. In contrast, linear operations, being invertible, may help preserve information. In other words, specific nonlinearities are a desirable property in a feature detector, but not necessarily in a neural relay. Accordingly, in as far as A1 neurons act as general purpose gateways to more specialized cortical areas<sup>10,11</sup>, their approximately linear response properties might help ensure an efficient transmission of information to these secondary areas.

As in most previous neurophysiological and behavioural studies of auditory localization, we determined the SRFs of A1 neurons using brief noise bursts corresponding to single, stationary sound sources. Examination of the responses of A1 neurons to other stimulus properties suggests that anatomically and functionally distinct regions may exist within this cortical field<sup>27,28</sup>. Consequently, our finding that A1 neurons analyse sound source location in a largely linear manner does not preclude the possibility that at least some of those neurons may respond to other stimulus features in a nonlinear way. For example, it has been reported that the



**Figure 4** Predicting changes in SRF structure for own-ear and foreign-ear stimuli. Spatial responses were mapped from VAS stimuli generated from the head and external ears of another ferret (foreign ears) instead of the animal's own ears. **a**, Own-ear and **b**, foreign-

ear observed and predicted SRFs for one single unit. Stimulus levels were as follows. VAS (dB above unit threshold): 35; FTRF-weighted VAS (dB SPL): 67; single tonal reverberant component (dB SPL): 65.

responses of some A1 cells to species-specific vocalizations<sup>29</sup> or to combinations of different sounds<sup>30</sup> may be well suited for processing certain attributes of natural sound sources. □

## Methods

### Spatial receptive field measurement

Extracellular recordings were performed in A1 of barbiturate anaesthetized ferrets, using standard techniques<sup>22</sup>. VAS stimuli for SRF recording were 20-ms gaussian noise bursts convolved with finite impulse response filters that mimicked the direction-dependent frequency transfer characteristics of the head and outer ears. The filters were based on acoustical recordings obtained from each individual animal before the electrophysiological recording sessions. Stimulus presentations from 224 virtual source directions were randomly interleaved. Responses were averaged over five stimulus presentations for each virtual direction. Continuous SRF maps were generated from the averaged responses by triangulation and smoothing with a 45° wide running average filter. SRF maps are shown here using a quartic-authalic equal area map projection. See ref. 22 for further details.

### Frequency-time response field measurement

Binaural FTRFs were determined by reverse correlation to random chord sequence stimuli similar to those introduced in ref. 16. Twenty-millisecond tone pips (5-ms rise/fall time) were started randomly at any of 60 frequency bands that spanned 0.5 to 32 kHz in a tenth of an octave steps. The probability of a tone pip starting in any one frequency band during any 5-ms time interval was set to values between 1 and 2%, so that, on average, 2.4 to 4.8 tone pips would be on simultaneously at any time during the random chord sequence. Tone onsets in different frequency bands and subsequent time intervals were statistically independent. Two statistically independent random chord sequences were presented simultaneously, one to each ear. Binaural FTRFs were determined by spike-triggered averaging of peri-spike random chord sequence segments. Binaural FTRFs were typically based on 500 to 1,500 spikes, which took between 15–20 min to collect. Determined in this way, FTRFs are usually interpreted as posterior probability distributions for the presence of a tonal stimulus in a particular frequency band and ear and at a particular time in the past, given that a spike was observed at time zero. To generate response predictions we use a converse interpretation, namely that the FTRF estimates a change in mean firing rate at a particular latency if acoustic energy is applied to a particular frequency band. Inherent in this interpretation is the assumption that the response of A1 neurons is adequately approximated as a sum of acoustic energies, weighted and delayed by frequency band according to the FTRF.

### Sound levels

Stimulus levels in these experiments were chosen arbitrarily, and varied from unit to unit. Our primary concern was to make good use of the available dynamic range of the digital signal processing equipment, and no attempt was made to match the sound levels of the VAS stimuli used to map the SRFs to the sound levels of the random chord sequences used to estimate the binaural FTRFs. In fact, matching the levels of these very different types of auditory stimuli would pose difficult technical problems. However, to allow a post hoc comparison between the sound levels used during the FTRF measurement and the SRF mapping we estimated the level of the VAS stimuli within the excitatory frequency bands revealed by the FTRF. These values are given as FTRF-weighted VAS levels (in dB relative to 20 µPa) in the figure legends, and give the average energy of the VAS stimuli weighted by a normalized frequency contour through the maximum of the binaural FTRF at the virtual sound direction where this value is greatest. Also stated is the approximate level of the VAS stimuli relative to the unit threshold (estimated by presenting VAS stimuli at varying attenuations) and the sound level of a single tonal component of the reverse correlation stimuli used to determine the binaural FTRF (the single tonal revcor component). In our data set, FTRF-weighted VAS levels varied from 0 to 90 dB SPL, while tonal revcor component levels varied from 30 to 70 dB SPL. VAS and revcor levels were uncorrelated, and over the range of levels tested, the quality of the linear predictions was independent of sound levels.

Received 25 May; accepted 7 September 2001.

- King, A. J., Schnupp, J. W. H. & Doubell, T. P. The shape of ears to come: dynamic coding of auditory space. *Trends Cogn. Sci.* **5**, 261–270 (2001).
- Jenkins, W. M. & Merzenich, M. M. Role of cat primary auditory cortex for sound-localization behavior. *J. Neurophysiol.* **52**, 819–847 (1984).
- Kavanagh, G. L. & Kelly, J. B. Contribution of auditory cortex to sound localization by the ferret (*Mustela putorius*). *J. Neurophysiol.* **57**, 1746–1766 (1987).
- Masterton, R. B. In *Acoustical Signal Processing in the Central Auditory System* (ed. Syka, J.) 1–17 (Plenum, New York, 1997).
- Hefner, H. E. & Hefner, R. S. Effect of bilateral auditory cortex lesions on sound localization in Japanese macaques. *J. Neurophysiol.* **64**, 915–931 (1990).
- Goldberg, J. M. & Brown, P. B. Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localisation. *J. Neurophysiol.* **32**, 613–636 (1969).
- Kitzes, L. M., Wrege, K. S. & Cassady, J. M. Patterns of responses of cortical cells to binaural stimulation. *J. Comp. Neurol.* **192**, 455–472 (1980).
- Yu, J. J. & Young, E. D. Linear and nonlinear pathways of spectral information transmission in the cochlear nucleus. *Proc. Natl Acad. Sci. USA* **97**, 11780–11786 (2000).
- Peña, J. L. & Konishi, M. Auditory spatial receptive fields created by multiplication. *Science* **292**, 249–252 (2001).
- Tian, B., Reser, D., Durham, A., Kustov, A. & Rauschecker, J. P. Functional specialization in rhesus monkey auditory cortex. *Science* **292**, 290–293 (2001).

- Recanzone, G. H., Guard, D. C., Phan, M. L. & Su, T. K. Correlation between the activity of single auditory cortical neurons and sound-localization behavior in the macaque monkey. *J. Neurophysiol.* **83**, 2723–2739 (2000).
- Wightman, F. L. & Kistler, D. J. In *Human Psychophysics* (eds Yost, W. A., Popper, A. N. & Fay, R. R.) 155–192 (Springer, New York, 1993).
- Hofman, P. M. & Van Opstal, A. J. Spectro-temporal factors in two-dimensional human sound localization. *J. Acoust. Soc. Am.* **103**, 2634–2648 (1998).
- Wightman, F. L. & Kistler, D. J. Monaural sound localization revisited. *J. Acoust. Soc. Am.* **101**, 1050–1063 (1997).
- Clarke, S., Bellmann, A., Meuli, R. A., Assal, G. & Steck, A. J. Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. *Neuropsychologia* **38**, 797–807 (2000).
- deCharms, R. C., Blake, D. T. & Merzenich, M. M. Optimizing sound features for cortical neurons. *Science* **280**, 1439–1443 (1998).
- Kowalski, N., Depireux, D. A. & Shamma, S. A. Analysis of dynamic spectra in ferret primary auditory cortex. II. Prediction of unit responses to arbitrary dynamic spectra. *J. Neurophysiol.* **76**, 3524–3534 (1996).
- Movshon, J. A., Thompson, I. D. & Tolhurst, D. J. Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *J. Physiol. Lond.* **283**, 53–77 (1978).
- Phillips, D. P., Judge, P. W. & Kelly, J. B. Primary auditory cortex in the ferret (*Mustela putorius*): neural response properties and topographic organization. *Brain Res.* **443**, 281–294 (1988).
- Carlike, S. The auditory periphery of the ferret. II: The spectral transformations of the external ear and their implications for sound localization. *J. Acoust. Soc. Am.* **88**, 2196–2204 (1990).
- Clarey, J. C., Barone, P. & Imig, T. J. In *The Mammalian Auditory Pathway: Neurophysiology* (eds Popper, A. & Fay, R.) 232–334 (Springer, New York, 1992).
- Mrsic-Flogel, T. D., King, A. J., Jenison, R. L. & Schnupp, J. W. H. Listening through different ears alters spatial response fields in ferret primary auditory cortex. *J. Neurophysiol.* **86**, 1043–1046 (2001).
- Brugge, J. F. et al. Simulation of free-field sound sources and its application to studies of cortical mechanisms of sound localization in the cat. *Hear. Res.* **73**, 67–84 (1994).
- Middlebrooks, J. C. & Pettigrew, J. D. Functional classes of neurons in primary auditory cortex of the cat distinguished by sensitivity to sound location. *J. Neurosci.* **1**, 107–120 (1981).
- Samson, F. K., Barone, P., Clarey, J. C. & Imig, T. J. Effects of ear plugging on single-unit azimuth sensitivity in cat primary auditory cortex. II. Azimuth tuning dependent upon binaural stimulation. *J. Neurophysiol.* **71**, 2194–2216 (1994).
- Ruggero, M. A. In *The Mammalian Auditory Pathway: Neurophysiology* (eds Popper, A. N. & Fay, R. R.) 34–93 (Springer, New York, 1992).
- Schreiner, C. E. Spatial distribution of responses to simple and complex sounds in the primary auditory cortex. *Audiol. Neurotol.* **3**, 104–122 (1998).
- Read, H. L., Winer, J. A. & Schreiner, C. E. Modular organization of intrinsic connections associated with spectral tuning in cat auditory cortex. *Proc. Natl Acad. Sci. USA* **98**, 8042–8047 (2001).
- Wang, X. On cortical coding of vocal communication sounds in primates. *Proc. Natl Acad. Sci. USA* **97**, 11843–11849 (2000).
- Nelken, I., Rotman, Y. & Bar Yosef, O. Responses of auditory-cortex neurons to structural features of natural sounds. *Nature* **397**, 154–157 (1999).

## Acknowledgements

We are grateful to D. Moore and A. Parker for comments on an earlier draft of this manuscript. This work was supported by Defeating Deafness (Dunhill Research Fellowship to J.W.H.S.) and by the Wellcome Trust (Wellcome Prize Studentship to T.D.M.F. and Wellcome Senior Research Fellowship to A.J.K.).

Correspondence and requests for materials should be addressed to J.W.H.S. (e-mail: jan.schnupp@physiol.ox.ac.uk).

# Target neuron prespecification in the olfactory map of *Drosophila*

Gregory S. X. E. Jefferis<sup>\*†‡</sup>, Elizabeth C. Marin<sup>†‡</sup>, Reinhard F. Stocker<sup>§</sup> & Liqun Luo<sup>\*†</sup>

<sup>\*</sup>Neurosciences Program, <sup>†</sup>Department of Biological Sciences, Stanford University, Stanford, California 94305, USA

<sup>§</sup>Department of Biology and Program in Neuroscience, University of Fribourg, CH-1700 Fribourg, Switzerland

<sup>‡</sup>These two authors contributed equally to this work

In *Drosophila* and mice, olfactory receptor neurons (ORNs) expressing the same receptors have convergent axonal projections to specific glomerular targets in the antennal lobe/olfactory bulb, creating an odour map in this first olfactory structure of the central nervous system<sup>1–3</sup>. Projection neurons of the *Drosophila*

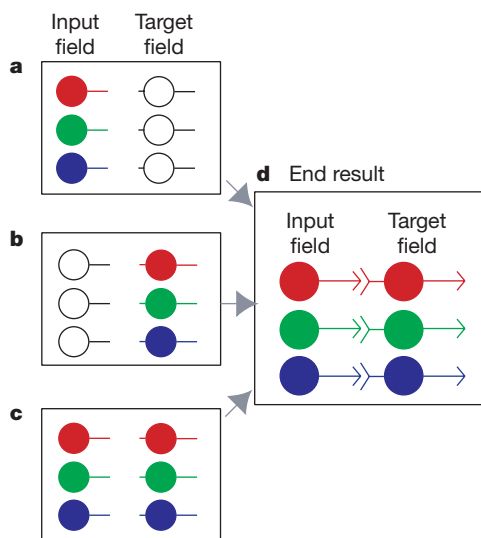


antennal lobe send dendrites into glomeruli and axons to higher brain centres<sup>4</sup>, thereby transferring this odour map further into the brain. Here we use the MARCM method<sup>5</sup> to perform a systematic clonal analysis of projection neurons, allowing us to correlate lineage and birth time of projection neurons with their glomerular choice. We demonstrate that projection neurons are prespecified by lineage and birth order to form a synapse with specific incoming ORN axons, and therefore to carry specific olfactory information. This prespecification could be used to hardwire the fly's olfactory system, enabling stereotyped behavioural responses to odorants. Developmental studies lead us to hypothesize that recognition molecules ensure reciprocally specific connections of ORNs and projection neurons. These studies also imply a previously unanticipated role for precise dendritic targeting by postsynaptic neurons in determining connection specificity.

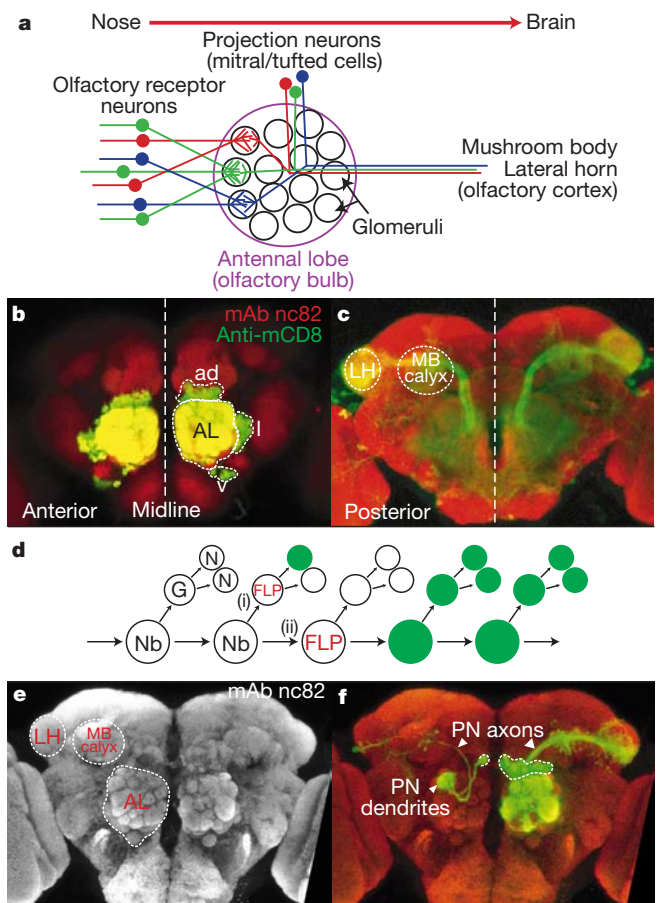
A common process in neural network formation is the establishment of one-to-one corresponding connections between two groups of neurons in two different locations, thereby generating a neural map (Fig. 1d). Three basic mechanisms for the formation of such neural maps can be proposed. In the first two mechanisms (Fig. 1a, b), either input or target neurons are genetically prespecified, whereas neurons of the remaining field are naive until specified by the identity of their partners during the connection process. In the third mechanism (Fig. 1c), input and target neurons are independently specified. Here, we explore this problem in the wiring of the *Drosophila* olfactory system.

The organization of the *Drosophila* peripheral olfactory pathway is very similar to that of mammals (Fig. 2a). About 1,300 ORNs expressing 40–60 different receptors<sup>2,6,7</sup> project their axons to 40–50 individually identifiable glomeruli of the antennal lobe<sup>8</sup> (equivalent to the mammalian olfactory bulb<sup>9</sup>). Information leaves the antennal lobe through an estimated 150 projection neurons (equivalent to mammalian mitral/tufted cells), whose cell bodies are located at the periphery of the antennal lobe<sup>4,10</sup> (Fig. 2b). Projection neurons

project their dendrites to glomeruli and their axons to higher brain centres, including the mushroom bodies and the lateral horn (Fig. 2a–c, see below). As in mice<sup>1</sup>, each ORN probably expresses one specific receptor<sup>2</sup>, and the axons of ORNs expressing the same receptors converge at the same morphologically and spatially distinct glomeruli<sup>2,3</sup>. In mice, ORNs seem to be genetically programmed to project to specific glomeruli, instructed by the receptors that they express<sup>11</sup>; indeed, this convergence seems to be independent of the presence of the target neurons of the olfactory bulb<sup>12</sup>. Although analogous experiments have not been reported in flies, ORNs expressing a particular receptor reside in stereotypic and discrete zones of the antennae and maxillary palps—the fly's olfactory appendages<sup>2,6,7</sup>. Assuming that ORN cell bodies do not relocate after their axons reach the antennal lobe, it seems probable that their glomerular targets are prespecified. Thus, of the three models for formation of the neural map (Fig. 1), the second



**Figure 1** Three mechanisms of establishing a neural map between two classes of neurons in different fields (d). **a**, Neurons in the input field are prespecified. Neurons in the target field acquire their identity from incoming neurons of the input field. **b**, Target neurons are prespecified. Input neurons acquire their identity from the target neurons with which they connect. **c**, Input and target neurons are independently specified. Matching colours signify that two neurons will eventually carry the same information, such as activation of a specific odorant receptor. Although we use one-to-one connections for simplicity, a similar logic applies to many-to-one (convergent) or one-to-many (divergent) connections.



**Figure 2** Organization of the antennal lobe. **a**, Schematic of the *Drosophila* olfactory system, with mammalian counterparts in parentheses. Particular colours represent ORNs expressing particular receptors and the projection neurons that form synapses with these ORNs. **b**, **c**, GAL4-GH146 drives marker (mouse CD8-GFP, green) expression in most of the projection neurons, counterstained with a general neuropil marker nc82 (red). Anterior confocal sections (**b**) show anterodorsal (ad), lateral (l) and ventral (v) cell-body clusters and their dendrites in the antennal lobe (AL); posterior sections (**c**) show axon tracts and their branches in the mushroom body (MB) and lateral horn (LH). mAb, monoclonal antibody. **d**, With MARCM<sup>5</sup> one can generate positively labelled single-cell clones (i) or neuroblast clones containing all neurons from a neuroblast lineage born after the mitotic recombination event (ii). Nb, neuroblast; G, ganglion mother cell; N, (postmitotic) neuron. **e**, **f**, Single-cell (left hemisphere) and neuroblast (right) projection neuron clones visualized by mouse CD8-GFP marker (green) counterstained with nc82 (red, shown alone in **e**). The cell bodies of each clone are surrounded by dotted lines. These and all subsequent images are anterior views with dorsal side up. PN, projection neuron.

model seems unlikely in both mice and *Drosophila*. We wanted to distinguish whether *Drosophila* projection neurons are specified by virtue of their connection with ORNs (Fig. 1a) or are independently specified (Fig. 1c).

The MARCM (mosaic analysis with a repressible cell marker) system<sup>5</sup> can be used to determine neuronal lineage and the projection patterns of individual neurons<sup>13</sup>. A typical neuroblast in the *Drosophila* brain undergoes asymmetric division to regenerate a new neuroblast and a ganglion mother cell, which divides once more to generate two postmitotic neurons. Using MARCM, one can generate positively-labelled single-cell clones as well as neuroblast clones (Fig. 2d). By controlling the timing of mitotic recombination using heat-shock-induced FLP recombinase, one can produce labelled clones of cells born at different developmental times. To study projection neurons of the antennal lobe, we made use of a GAL4 line, GAL4-GH146 (ref. 14), which drives marker expression in a large subset of projection neurons (approximately 90). By crossing GH146 with a membrane marker, UAS-mouse CD8-green fluorescent protein (GFP)<sup>5</sup>, we can visualize the cell body and dendrites of projection neurons in the antennal lobe in the anterior part of the brain (Fig. 2b), as well as their axonal projections in the posterior part of the brain (Fig. 2c). When we apply the MARCM technique using GAL4-GH146, we can selectively visualize subsets of these projection neurons as either neuroblast clones or single-cell clones (Fig. 2f).

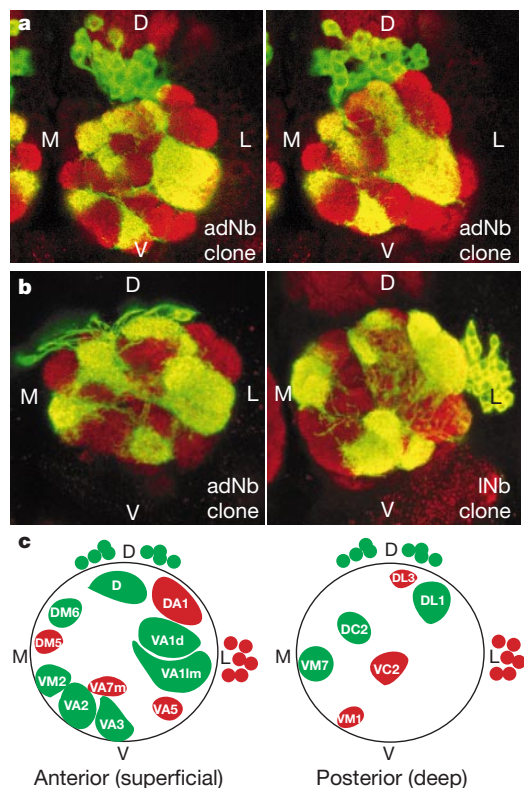
Systematic clonal analysis revealed that GH146-positive projection neurons (to which we refer hereafter as projection neurons) are derived from three neuroblasts: an anterodorsal, a lateral and a

ventral neuroblast, corresponding to the three groups of cell body outlined in Fig. 2b. When neuroblast clones are induced in the early embryo and then examined in adults, the anterodorsal, lateral and ventral neuroblasts give rise to approximately 50, 35 and 6 projection neurons, respectively. These three numbers correspond well to the number of projection neurons present in the three GH146-positive cell groups (Fig. 2b). As we cannot induce any clones by applying heat shock after puparium formation (APF), and neuroblast clones generated in late larvae contain 3–5 cells, we infer that all projection neurons are born well before the arrival of pioneering adult ORN axons in the antennal lobe around 20–24 h APF<sup>15</sup>. We focused our subsequent analyses on anterodorsal projection neurons and lateral projection neurons as most of these neurons have uniglomerular dendritic projections (see Fig. 2f, Fig. 4b), whereas some ventral projection neurons have diffuse dendritic arborizations, and all project by means of a different path to the higher brain centres, bypassing the mushroom bodies (our own unpublished data).

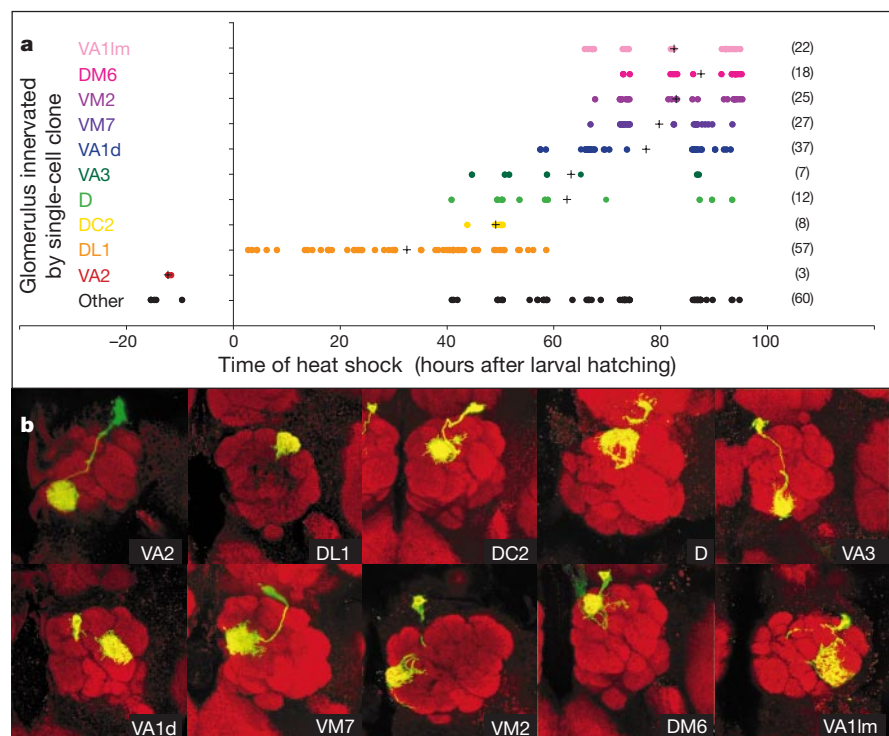
When glomerular projections of neuroblast clones generated in early larvae were examined, we found that anterodorsal and lateral neuroblast clones appear to innervate stereotypical, intercalated but non-overlapping glomeruli. Figure 3a shows a single confocal section of two independently generated anterodorsal neuroblast clones, revealing almost identical glomerular projections. By contrast, an anterodorsal neuroblast clone and a lateral neuroblast clone examined at the same depth project to complementary glomeruli (Fig. 3b). Figure 3c summarizes the 'landmark' glomeruli (see Methods) that are innervated by anterodorsal projection neurons (green) or lateral projection neurons (red). In the 54 anterodorsal neuroblast and 25 lateral neuroblast clones we examined, we found no exception to this rule, despite the fact that there is no obvious relationship between the positions of projection neuron cell bodies and their glomeruli<sup>4</sup>. These observations indicate that the neuroblast from which a projection neuron is derived restricts its glomerular choice and consequently the subset of olfactory information that it can carry further into the brain.

We next asked whether projection neurons are further specified within a neuroblast lineage. Because labelled, single-cell MARCM clones are born shortly after heat-shock induction of mitotic recombination<sup>13</sup>, we could test whether projection neurons born during specific developmental periods would project to specific glomeruli by using the time of heat shock as a variable (see Methods). We found that anterodorsal single-cell projection neuron clones with particular glomerular projections were generated within restricted and characteristic developmental windows (Fig. 4a, b). Notably, single-cell clones induced by early larval heat shock (0–36 h) exclusively produced projection neurons projecting to glomerulus DL1.

Because individual larvae develop at different rates, and FLP recombinase can persist for some time after heat-shock induction, single-cell clone analysis cannot distinguish unequivocally the birth order of projection neurons that are born immediately after each other (Fig. 4a). We therefore took a complementary approach in which we examined multicellular neuroblast clones generated at different developmental periods, and determined whether they included certain landmark glomeruli. If projection neurons innervating different glomeruli are produced in a defined sequence, then multicellular neuroblast clones (Fig. 2d) induced at progressively later times during development should innervate a subset of the glomeruli in neuroblast clones generated at earlier times. Eventually the last-born, smallest clones should contain projection neurons innervating only one glomerulus. Such a nested set is exactly what we observed when anterodorsal neuroblast clones were scored for the presence or absence of projection neurons innervating ten landmark glomeruli (Fig. 5). We could therefore infer an ordered birth sequence of projection neurons (VA2, DL1, DC2, D, VA3, VA1d, VM7, VM2, DM6, VA1lm) from the 54 clones analysed. We



**Figure 3** Dorsal and lateral neuroblast clones contain projection neurons with stereotypical and complementary glomerular projections. **a**, Single, confocal sections of two independent anterodorsal neuroblast (adNb) clones taken at similar depths. **b**, Single, confocal sections of an anterodorsal neuroblast and a lateral neuroblast (lNb) clone taken at similar depths. **c**, Schematic of landmark glomeruli derived from anterodorsal neuroblast (green) and lateral neuroblast (red) in the anterior (superficial) or posterior (deep) sections of the antennal lobe. D, dorsal; V, ventral; M, medial; L, lateral.

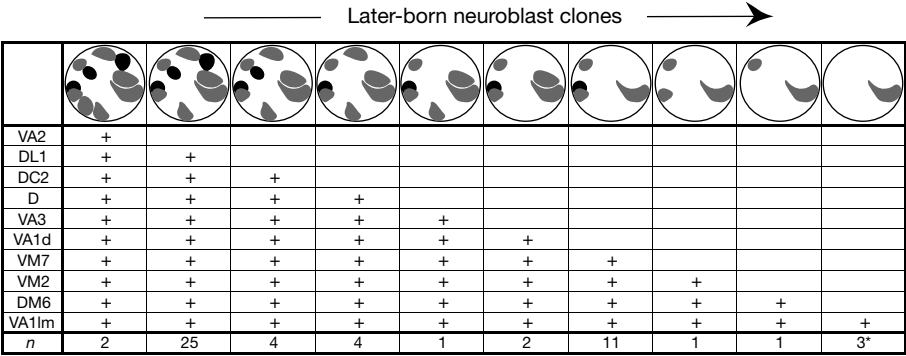


**Figure 4** Single-cell clone analysis. **a**, Glomerular identity plotted against time of heat-shock induction for 276 single-cell clones from the anterodorsal neuroblast lineage. Each dot represents a single clone, and crosses represent the mean heat-shock time for a

particular glomerular class. Numbers in parentheses represent the total single-cell clones per glomerular class. **b**, Representative images of the ten landmark single-cell clone classes.

did not find a neuroblast clone in which projection neurons projecting to a particular glomerulus were altered from this order; for example, neuroblast clones containing projection neurons projecting to VM7 always additionally contained all three of the later-born types of projection neuron, VM2, DM6 and VA1Im. As, on average, about three projection neurons innervate each glomerulus, the fact that we can infer an order implies that projection neurons innervating a common glomerulus are likely to be born at a similar time. Together with the single-cell clone analysis (Fig. 4), we conclude that, at least for these ten glomeruli, there is a strict order of generation of projection neurons that can predict future glomerular targets.

How can the birth time of a projection neuron predict which glomerulus it will eventually innervate? One possibility is that the ordered generation of projection neurons could result in the ordered differentiation of their dendrites, and that temporally ordered availability of proto-glomeruli for innervation restricts projection neurons born at a certain time to a particular glomerulus. However, we found that at around 22 h APF, when pioneering ORN axons just start to invade the antennal lobe<sup>15</sup>, projection neurons born at different times had similar dendritic differentiation statuses, having already initiated their dendritic branches in the vicinity of the antennal lobe region. The axons of the projection neurons had already reached the mushroom body and the lateral



**Figure 5** Ordered generation of projection neurons derived from neuroblast clone analysis. Anterodorsal neuroblast clones were scored for the presence (+) or absence of landmark glomeruli; all 54 clones fell into one of ten classes (top). Although we did not use the time of clone induction to deduce order, larger clones were generated from earlier

heat shocks (as represented by the arrow) with few exceptions. Darker glomeruli represent those in deeper sections. Asterisk, these clones are neuroblast rather than single-cell clones because they contain 3–5 projection neuron cell bodies.



horn (see Supplementary Information Fig. 1). These observations argue against the hypothesis of differentiation timing. We instead favour the hypothesis that individual projection neurons and ORNs are independently specified to carry molecular signals that allow them to recognize either each other or a common set of cues located at the developing antennal lobe (Fig. 1c). Although it is assumed that stereotyped ORN axon projections<sup>2,3</sup> depend on a guidance map in the developing antennal lobe, our results imply that such cues may also be used for precise dendritic targeting of projection neurons, or that projection neuron dendrites actively participate in creating the guidance map.

Independent pre-patterning of input and target fields has been demonstrated in the formation of vertebrate retinotectal projections along the anterior–posterior axis<sup>16,17</sup>, and even implicated in the development of ocular dominance columns<sup>18</sup>. In both systems, activity-dependent processes have important roles in refining the coarse map at the level of the single cell<sup>19,20</sup>. We show here the independent specification of projection neurons at the single-cell level—matching the precision of the ORN identities. These experiments support the importance of independent specification (Fig. 1c) in formation of the neural map. Moreover, as no obvious logic correlates cell body position of the projection neuron, birth time and the location of the glomerular projection, simple molecular gradient/counter-gradient models as used for axon guidance in the retinotectal system<sup>16,17</sup> are unlikely to suffice. Instead, we propose that dendrites of projection neurons use a combination of recognition molecules specific to each eventual glomerulus; mechanisms that generate a complex repertoire of cell-surface molecules have recently been described<sup>21,22</sup>. Furthermore, our study uncovers an elegant mechanism for specifying different projection neurons: the precisely ordered generation by a single neuroblast of a large number of distinct neurons. We are in the process of determining whether a timer mechanism intrinsic to the neuroblast, cues from neighbouring cells at the time of birth, or a combination of such mechanisms<sup>23–25</sup> are used, perhaps to specify the expression of recognition molecules.

Given the similarities in organization of the *Drosophila* and mammalian peripheral olfactory systems, it will be of great interest to test whether and to what extent the mitral/tufted cells in the mammalian olfactory system are independently specified. It is conceivable that the information used to pattern the olfactory bulb for ORN axon targeting could be used to prespecify mitral/tufted cells, thereby coordinating their dendritic targets in the olfactory bulb and axonal projections in higher brain centres. Such prespecification mechanisms may also be used in neural map formation in other parts of the developing brain. □

## Methods

### Clonal analysis

In the MARCM strategy, a cell marker is under the control of a GAL4-UAS promoter, which is activated by GAL4 and repressed by GAL80 (ref. 5). The GAL80 transgene is driven by a ubiquitous promoter and placed distal to an FLP-mediated recombination site. Flies heterozygous for the GAL80 transgene and carrying appropriate GAL4 and UAS-marker transgenes do not normally express the marker owing to inhibition by GAL80. Only after FLP-mediated mitotic recombination will the marker be expressed in cells that have lost the GAL80 transgene and express the GAL4 transgene. We maintained flies on standard medium at 25 °C. Larvae of the genotype *y w hs-FLP UAS-mCD8-GFP/+ or Y; FRT<sup>G13</sup> tubP-GAL80/FRT<sup>G13</sup> GAL4-GH146 UAS-mCD8-GFP* were collected over a 2-h period. After appropriate aging, they were given a 1-h heat shock at 37 °C. To verify that clones homozygous for the FRT<sup>G13</sup> and GAL4-GH146 insertions (both located on chromosome arm 2R) exhibit no abnormalities, we performed control experiments with larvae of the genotype *y w hs-FLP UAS-mCD8-GFP/Y; tubP-GAL80 FRT<sup>40A</sup>/FRT<sup>40A</sup> GAL4-GH146 UAS-mCD8-GFP*, and observed no difference in clone composition or neuronal structure.

### Immunocytochemistry and imaging

Adult brains were dissected out at least 48-h after eclosion. Fixation and immunocytochemistry were carried out as described<sup>13</sup> using the following antibodies: rat monoclonal anti-mouse CD8  $\alpha$ -subunit (Caltag), 1:100; mouse monoclonal nc82 (a gift of E. Buchner and

A. Hofbauer), 1:20; mouse monoclonal 1D4 (gift of C. Goodman), 1:20; Alexa-488 conjugated goat anti-rat immunoglobulin- $\gamma$  (IgG), 1:200; and Alexa-568 conjugated goat anti-mouse IgG, 1:200 (Molecular Probes). Stacks of optical sections, usually at 0.5- $\mu$ m spacing, were obtained with a Bio-Rad MRC 1024 laser-scanning confocal microscope, using the LaserSharp image-collection program, then processed with NIH Image and Adobe Photoshop.

### Glomerular analysis

A subset of the previously recognized glomeruli<sup>8</sup> were selected as landmark glomeruli because we were unambiguously able to recognize them in neuroblast clones. In the anterodorsal neuroblast lineage, projection neurons born in embryonic and larval stages innervate approximately 5 and 13 glomeruli, respectively. For the lateral neuroblast all GH146-positive projection neurons are probably born later than 24 h after larval hatching and innervate approximately 12 glomeruli. We were unable to generate sufficient partial lateral neuroblast clones to determine the order of lateral projection neuron generation, so we restricted the birth order analysis to the anterodorsal neuroblast. Figures 3–5 tabulate information from more than 4,000 brains examined.

Received 3 July; accepted 27 September 2001.

1. Axel, R. The molecular logic of smell. *Sci. Am.* **273**, 154–159 (1995).
2. Vosshall, L. B., Wong, A. M. & Axel, R. An olfactory sensory map in the fly brain. *Cell* **102**, 147–159 (2000).
3. Gao, Q., Yuan, B. & Chess, A. Convergent projections of *Drosophila* olfactory neurons to specific glomeruli in the antennal lobe. *Nature Neurosci.* **3**, 780–785 (2000).
4. Stocker, R. F., Lienhard, M. C., Borst, A. & Fischbach, K.-F. Neuronal architecture of the antennal lobe in *Drosophila melanogaster*. *Cell Tissue Res.* **262**, 9–34 (1990).
5. Lee, T. & Luo, L. Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron* **22**, 451–461 (1999).
6. Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky, A. & Axel, R. A spatial map of olfactory receptor expression in the *Drosophila* antenna. *Cell* **96**, 725–736 (1999).
7. Clyne, P. J. et al. A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*. *Neuron* **22**, 327–338 (1999).
8. Laissue, P. P. et al. Three-dimensional reconstruction of the antennal lobe in *Drosophila melanogaster*. *J. Comp. Neurol.* **405**, 543–552 (1999).
9. Hildebrand, J. G. & Shepherd, G. M. Mechanisms of olfactory discrimination: converging evidence for common principles across phyla. *Annu. Rev. Neurosci.* **20**, 595–631 (1997).
10. Stocker, R. F. The organization of the chemosensory system in *Drosophila melanogaster*: a review. *Cell Tissue Res.* **275**, 3–26 (1994).
11. Mombaerts, P. Molecular biology of odorant receptors in vertebrates. *Annu. Rev. Neurosci.* **22**, 487–509 (1999).
12. Bulfone, A. et al. An olfactory sensory map develops in the absence of normal projection neurons or GABAergic interneurons. *Neuron* **21**, 1273–1282 (1998).
13. Lee, T., Lee, A. & Luo, L. Development of the *Drosophila* mushroom bodies: sequential generation of three distinct types of neurons from a neuroblast. *Development* **126**, 4065–4076 (1999).
14. Stocker, R. F., Heimbeck, G., Gendre, N. & de Belle, J. S. Neuroblast ablation in *Drosophila* P[GAL4] lines reveals origins of olfactory interneurons. *J. Neurobiol.* **32**, 443–452 (1997).
15. Jhaveri, D., Sen, A. & Rodrigues, V. Mechanisms underlying olfactory neuronal connectivity in *Drosophila*—the atonal lineage organizes the periphery while sensory neurons and glia pattern the olfactory lobe. *Dev. Biol.* **226**, 73–87 (2000).
16. Sperry, R. W. Chemoaffinity in the orderly growth of nerve fiber patterns and connections. *Proc. Natl Acad. Sci. USA* **50**, 703–710 (1963).
17. Flanagan, J. G. & Vanderhaeghen, P. The ephrins and Eph receptors in neural development. *Annu. Rev. Neurosci.* **21**, 309–345 (1998).
18. Crowley, J. C. & Katz, L. C. Early development of ocular dominance columns. *Science* **290**, 1321–1324 (2000).
19. Goodman, C. S. & Shatz, C. J. Developmental mechanisms that generate precise patterns of neuronal connectivity. *Cell* **72**, 77–98 (1993).
20. Katz, L. C. & Shatz, C. J. Synaptic activity and the construction of cortical circuits. *Science* **274**, 1133–1138 (1996).
21. Wu, Q. & Maniatis, T. A striking organization of a large family of human neural cadherin-like cell adhesion genes. *Cell* **97**, 779–790 (1999).
22. Schmucker, D. et al. Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* **101**, 671–684 (2000).
23. McConnell, S. K. Strategies for the generation of neuronal diversity in the developing central nervous system. *J. Neurosci.* **15**, 6987–6998 (1995).
24. Edlund, T. & Jessell, T. M. Progression from extrinsic to intrinsic signaling in cell fate specification: a view from the nervous system. *Cell* **96**, 211–224 (1999).
25. Ishiki, T., Pearson, B., Holbrook, S. & Doe, C. Q. *Drosophila* neuroblasts sequentially express transcription factors which specify the temporal identity of their neuronal progeny. *Cell* **106**, 511–521 (2001).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

### Acknowledgements

We thank members of the Luo laboratory and B. Baker, S. McConnell, W. Newsome, J. Ngai, C. Niell and D. Hoepfner for comments on the manuscript. G.S.X.E.J. and E.C.M. are HHMI pre-doctoral fellows. This work was supported by a Swiss National Funds grant to R.F.S. and NIH grants and a Terman Fellowship to L.L.

Correspondence and requests for materials should be addressed to L.L. (e-mail: lluo@stanford.edu).

# An anorexic lipid mediator regulated by feeding

F. Rodríguez de Fonseca<sup>\*,†</sup>, M. Navarro<sup>\*</sup>, R. Gómez<sup>\*</sup>, L. Escuredo<sup>\*</sup>, F. Nava<sup>‡</sup>, J. Fu<sup>‡</sup>, E. Murillo-Rodríguez<sup>‡</sup>, A. Giuffrida<sup>‡</sup>, J. LoVerme<sup>‡</sup>, S. Gaetani<sup>‡</sup>, S. Kathuria<sup>‡</sup>, C. Gall<sup>||</sup> & D. Piomelli<sup>‡</sup>

<sup>\*</sup> Department of Psychobiology, Complutense University, 28233 Madrid, Spain

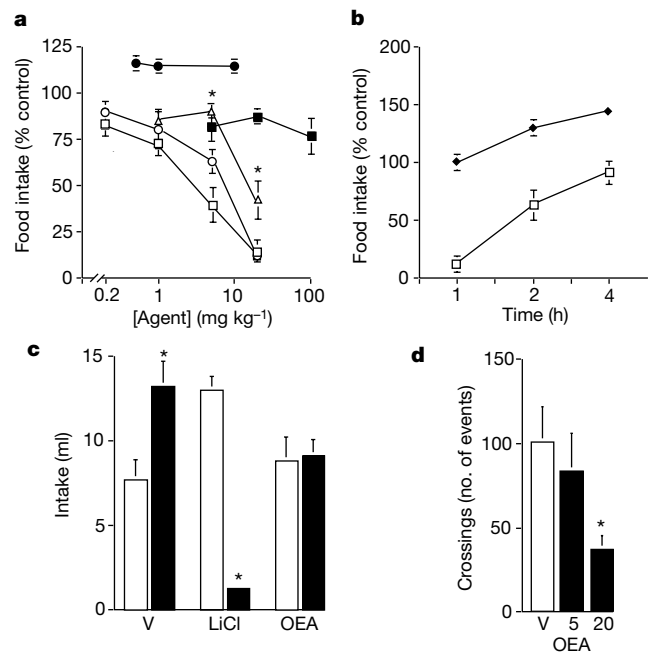
<sup>†</sup> Fundación Hospital Carlos Haya, 29010 Málaga, Spain

<sup>‡</sup> Department of Pharmacology, and <sup>||</sup> Department of Anatomy and Neurobiology, University of California, Irvine, California 92697, USA

Oleylethanolamide (OEA) is a natural analogue of the endogenous cannabinoid anandamide. Like anandamide, OEA is produced in cells in a stimulus-dependent manner and is rapidly eliminated by enzymatic hydrolysis, suggesting a function in cellular signalling<sup>1</sup>. However, OEA does not activate cannabinoid receptors and its biological functions are still unknown<sup>2</sup>. Here we show that, in rats, food deprivation markedly reduces OEA biosynthesis in the small intestine. Administration of OEA causes a potent and persistent decrease in food intake and gain in body mass. This anorexic effect is behaviourally selective and is associated with the discrete activation of brain regions (the paraventricular hypothalamic nucleus and the nucleus of the solitary tract) involved in the control of satiety. OEA does not affect food intake when injected into the brain ventricles, and its anorexic actions are prevented when peripheral sensory fibres are removed by treatment with capsaicin. These results indicate that OEA is a lipid mediator involved in the peripheral regulation of feeding.

Fatty acid ethanolamides (FAEs) are unusual components of animal and plant lipids<sup>3–5</sup> that are synthesized in response to a variety of physiological and pathological stimuli, including activation of neurotransmitter receptors in rat brain neurons<sup>1,6</sup> and exposure to metabolic stressors in mouse epidermal cells<sup>7</sup>. The primary mechanism underlying FAE generation in mammalian tissues involves two concerted biochemical reactions: cleavage of the membrane phospholipid *N*-acyl phosphatidylethanolamine (NAPE), catalysed by an unknown phospholipase D; and NAPE re-synthesis, mediated by an *N*-acyltransferase (NAT) that is regulated by calcium ions and cyclic AMP<sup>8,9</sup>. After release, FAEs are transported back into cells<sup>10</sup> and eventually broken down to fatty acid and ethanolamine by an intracellular fatty acid amide hydrolase (FAAH)<sup>11,12</sup>. That animal cells release FAEs in a stimulus-dependent manner suggests that these compounds may participate in cell-to-cell communication. Further support for this idea comes from the discovery that the polyunsaturated FAE anandamide (arachidonyl-ethanolamide) serves as an endogenous ligand for cannabinoid receptors<sup>13</sup>. However, the pharmacological effects of saturated or mono-unsaturated FAEs such as OEA cannot be accounted for by activation of any of the known cannabinoid receptor subtypes<sup>2</sup>, and the biological roles of these compounds remain elusive.

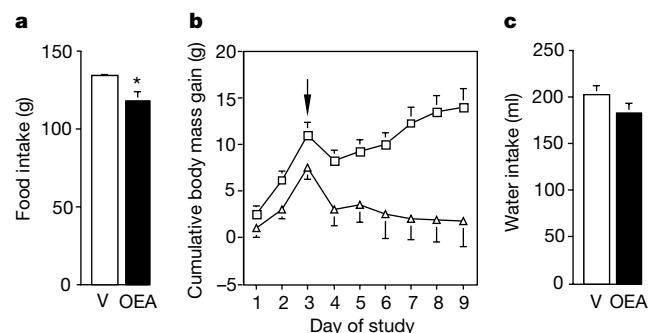
Because anandamide may regulate feeding<sup>14</sup>, we investigated the effects of OEA on food intake in rats. Systemic administration of OEA caused a dose- and time-dependent suppression of food consumption (Fig. 1). Under the same conditions anandamide and oleic acid had no effect, palmitylethanolamide was significantly less potent than OEA and elaidylethanolamide was similar in potency to OEA (Fig. 1a). These results indicate that OEA reduces eating in a structurally selective manner, and suggest that the molecular requisites for this effect are distinct from those involved in the interaction of anandamide with recognized cannabinoid



**Figure 1** OEA suppresses food intake in rats deprived of food for 24 h. **a**, Dose-dependent effects of i.p. OEA (open squares), elaidylethanolamide (open circles), palmitylethanolamide (triangles), oleic acid (filled squares) and anandamide (filled circles) after the first 60 min of food presentation. Food intake after injection of vehicle (70% DMSO in saline, 1 ml kg<sup>-1</sup> i.p.) was 7.1 ± 0.5 g per animal (100%). **b**, Time course of the hypophagic effects of i.p. OEA (20 mg kg<sup>-1</sup>) (squares) or vehicle (diamonds) on food intake. **c**, Effects of i.p. vehicle (V), lithium chloride (LiCl; 0.4 M, 7.5 ml kg<sup>-1</sup>) or OEA (20 mg kg<sup>-1</sup>) on conditioned taste aversion. Open bars, water intake; solid bars, saccharin intake. **d**, Effects of i.p. vehicle (V) or OEA (5 or 20 mg kg<sup>-1</sup>) on horizontal activity in an open field, assessed on habituated animals. Asterisk, *P* < 0.05; *n* = 8–12 per group.

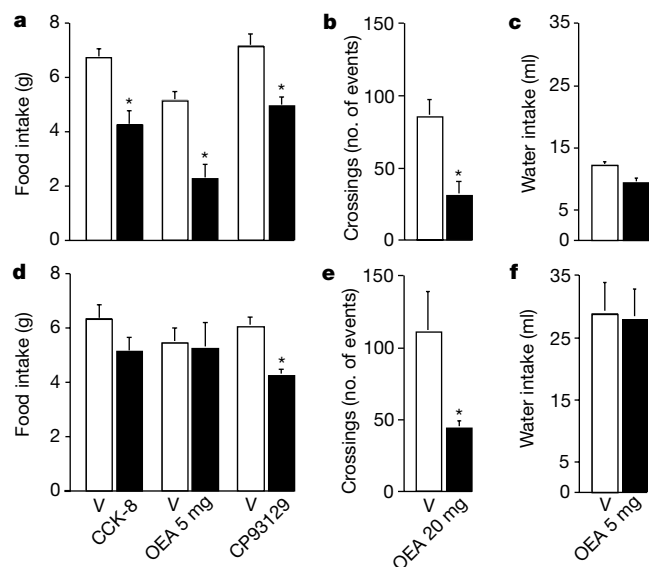
targets<sup>15</sup>. In further support of this idea, CB1 and CB2 cannabinoid antagonists (SR141716A and SR144528, respectively) did not affect OEA hypophagia (data not shown).

It is possible that OEA reduced food intake by inducing a non-specific state of behavioural suppression. If this were the case, OEA should cause conditioned taste aversion, which can be readily provoked in rats by administration of lithium chloride (Fig. 1c). By contrast, a high dose of OEA had little effect in this assay, suggesting that the compound is not aversive (Fig. 1c). Furthermore, OEA did not produce anxiety-like symptoms, did not affect



**Figure 2** Effects of subchronic OEA administration on food intake and body mass. **a**, Effects of i.p. vehicle (V; 5% Tween 80 with 5% propylene glycol in sterile saline; open bars) or OEA (5 or 20 mg kg<sup>-1</sup>, once per day; solid bars) on cumulative food intake. **b**, Time course of the effects of OEA (triangles) or vehicle (squares) on body mass gain. The arrow marks the administration of vehicle or OEA. **c**, Effects of vehicle or OEA on cumulative water intake. Asterisk, *P* < 0.05; *n* = 10 per group.

§ Present addresses: Department of Neuroscience 'B. B. Brodie', University of Cagliari, Cagliari, Italy (F.N.); Department of Physiology, Universidad Nacional Autónoma de México, Mexico City, Mexico (E.M.-R.).



**Figure 3** Role of peripheral sensory fibres in OEA-induced hypophagia. Effects of i.p. vehicle (V), OEA (5 mg kg<sup>-1</sup>), CCK-8 (10 µg kg<sup>-1</sup>) and CP-93129 (1 mg kg<sup>-1</sup>), a centrally active 5-HT<sub>1B</sub> receptor agonist, on food intake in control rats (**a–c**) and capsaicin-treated rats (**d–f**). **a, d**, Food intake. **b, e**, Horizontal activity in an open field. **c, f**, Water intake. Asterisk,  $P < 0.05$ ;  $n = 8–12$  per group.

activity of the hypothalamus–pituitary–adrenal (HPA) axis, did not change body temperature, pain threshold, plasma glucose levels (vehicle:  $1.11 \pm 0.04$  mg ml<sup>-1</sup>; OEA (5 mg kg<sup>-1</sup> intraperitoneally, i.p.): 30 min,  $1.10 \pm 0.03$  mg ml<sup>-1</sup>; 60 min,  $0.99 \pm 0.02$  mg ml<sup>-1</sup>; 120 min,  $1.08 \pm 0.07$  mg ml<sup>-1</sup>), insulin levels (vehicle:  $1.7 \pm 0.3$  ng ml<sup>-1</sup>; OEA (5 mg kg<sup>-1</sup> i.p.): 30 min,  $1.5 \pm 0.1$  ng ml<sup>-1</sup>; 60 min,  $1.6 \pm 0.1$  ng ml<sup>-1</sup>; 120 min,  $1.6 \pm 0.2$  ng ml<sup>-1</sup>;  $n = 5$  or 6) or leptin levels (vehicle:  $13.3 \pm 2.4$  ng ml<sup>-1</sup>; OEA (5 mg kg<sup>-1</sup> i.p.): 30 min,  $10.5 \pm 0.4$  ng ml<sup>-1</sup>; 60 min,  $11.4 \pm 1.1$  ng ml<sup>-1</sup>; 120 min,  $11.6 \pm 0.8$  ng ml<sup>-1</sup>;  $n = 5$ ) (see also Supplementary Information). Although a high OEA dose reduced motor activity (Fig. 1d), this effect may not contribute to the compound's anorectic actions, for two reasons: the same dose of OEA did not affect water intake (see Supplementary Information); and removal of sensory fibres by capsaicin treatment abrogated the effects of OEA on food intake, but not those on movement (see below). This pharmacological profile differentiates OEA from other appetite suppressants such as amphetamine and glucagon-like peptide 1 (the effects of which include aversion and anxiety), and from cannabinoids such as anandamide (which stimulates food intake in pre-satiated animals, reduces pain and activates the HPA axis)<sup>16</sup>.

To test whether tolerance develops to the hypophagic actions of OEA, we administered it subchronically in rats. Daily injections of OEA (5 mg kg<sup>-1</sup> i.p.) resulted in a small but significant decrease in cumulative food intake (Fig. 2a), which was accompanied by a marked inhibition of body mass gain (Fig. 2b). By contrast, subchronic OEA administration had no effect on water intake (Fig. 2c) or on plasma levels of various metabolites and liver enzymes (see Supplementary Information). Pair-feeding experi-

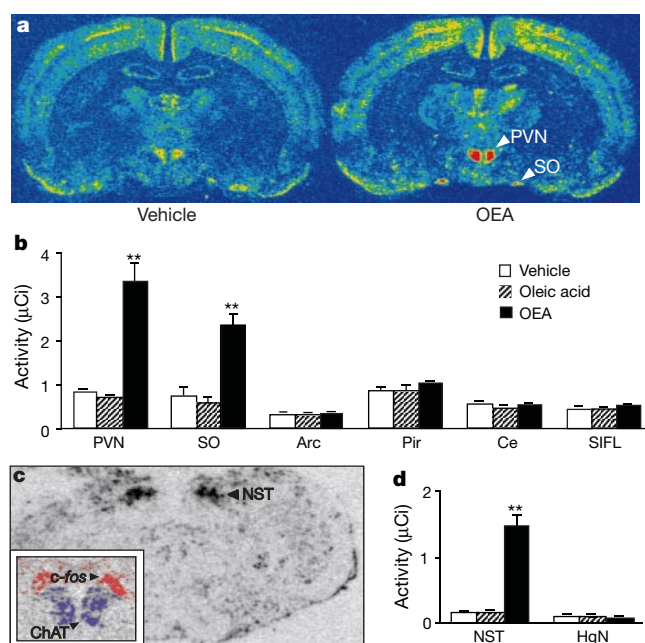
**Table 1** Effects of OEA and paired feeding on body mass gain

|          | <i>n</i> | Cumulative mass gain (g) | Food intake (g) | Water intake (ml) |
|----------|----------|--------------------------|-----------------|-------------------|
| Control  | 10       | 15.6 ± 2.3               | 143.2 ± 10.3    | 184.3 ± 12.1      |
| OEA      | 11       | 4.1 ± 1.4†               | 115.9 ± 4.2*    | 174.7 ± 9.1       |
| Pair fed | 11       | 4.3 ± 1.4†               | 116             | 192.8 ± 12.7      |

OEA (5 mg kg<sup>-1</sup> i.p.) or vehicle (5% propylglycol, 5% Tween 80, 90% sterile saline) were administered once per day for 7 days, after 3 days of baseline measurements. The pair-fed group received a total of 116 g of food, identical to the amount consumed by the OEA group.

\*  $P < 0.05$  versus control.

†  $P < 0.01$  versus control.



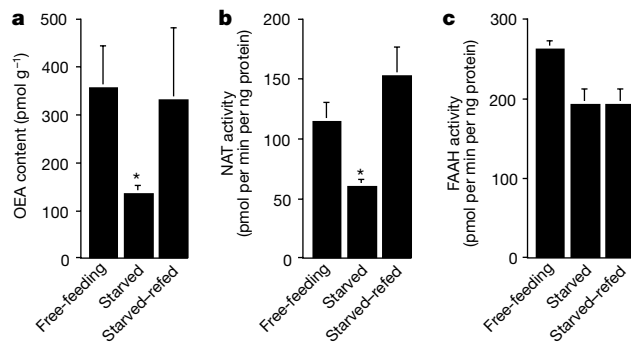
**Figure 4** OEA increases *c-fos* mRNA expression in brain regions associated with feeding behaviour. **a**, Pseudocolour images of autoradiograms showing that i.p. OEA (right) elicits a profound and selective increase in *c-fos* mRNA expression in the paraventricular (PVN) and supraoptic (SO) hypothalamic nuclei, as assessed by *in situ* hybridization. A representative section from a vehicle-treated rat is shown (left). Labelling densities are indicated by colour: (low) blue–green–yellow–red (high). **b**, Quantification of *c-fos* mRNA labelling in forebrain regions (PVN, SO, arcuate (Arc), layer II piriform cortex (Pir), central nucleus of the amygdala (Ce) and S1 forelimb cortex (S1FL)) of rats treated with i.p. vehicle, OEA or oleic acid. **c**, Autoradiogram showing elevated *c-fos* mRNA expression in the NST in an OEA-treated rat. The inset shows *c-fos* antisense RNA labelling in the NST (red) identified by its localization relative to adjacent efferent nuclei (hypoglossal and dorsal motor nucleus of the vagus), which express choline acetyl transferase (ChAT) mRNA (purple). **d**, OEA increases *c-fos* mRNA expression in the NST, but not in the hypoglossal nucleus (HgN). Double asterisk,  $P < 0.0001$ ;  $n = 5$  per group.

ments suggest that decreased food consumption is sufficient to account for the mass-reducing actions of OEA (Table 1). We cannot exclude, however, the possible contribution of other factors to this effect, such as stimulation of energy expenditure or inhibition of energy accumulation, as suggested by the significantly lower triglyceride levels in OEA-treated rats than in controls (see Supplementary Information).

Although potent when administered peripherally, OEA is ineffective after intracerebroventricular injection (see Supplementary Information), leading us to hypothesize that its primary sites of action are located outside the central nervous system (CNS). To test this idea, we destroyed peripheral sensory fibres by treating adult rats with capsaicin<sup>17</sup>. Capsaicin-treated rats failed to respond to systemically administered cholecystokinin-8 (CCK-8) (Fig. 3a, d) and drank more water than controls (Fig. 3c, f), two indications that the neurotoxin had removed sensory afferents<sup>18</sup>. Capsaicin-treated animals also failed to become hypophagic in response to OEA (5 mg kg<sup>-1</sup> i.p.), but responded normally to the compound CP-93129, which targets brain 5-HT<sub>1B</sub> receptors (Fig. 3a, d)<sup>19</sup>. These findings support the hypothesis that OEA reduces food intake by acting at a peripheral site and that sensory fibres are required for this effect. Interestingly, the ability of a high OEA dose (20 mg kg<sup>-1</sup> i.p.) to reduce motor activity was not affected by capsaicin treatment (Fig. 3b, e), suggesting that OEA modulates eating and movement through distinct mechanisms.

Peripheral inputs related to appetite suppression recruit CNS structures such as the nucleus of the solitary tract (NST) in the





**Figure 5** Feeding regulates OEA biosynthesis in small intestine. Effects of free feeding, starvation and starvation–refeeding on OEA levels (**a**), NAT activity levels (**b**) and FAAH activity levels (**c**). Asterisk,  $P < 0.05$ ;  $n = 3$ –5 per group.

brainstem and the paraventricular (PVN) nucleus in the hypothalamus<sup>20</sup>. To identify brain pathways engaged during OEA-evoked hypophagia, we mapped messenger RNA levels for the activity regulated gene *c-fos*<sup>21</sup> by *in situ* hybridization after systemic administration of OEA, oleic acid or vehicle. When compared with controls, OEA (10 mg kg<sup>-1</sup>, i.p.) evoked a highly localized increase in *c-fos* mRNA levels in the PVN, supraoptic nucleus (Fig. 4a) and NST (Fig. 4c). This enhancement was specific, insofar as *c-fos* expression in other brain regions was not significantly affected by OEA (Fig. 4b, d). The stimulation of *c-fos* expression in the NST (which processes vagal sensory inputs to the CNS) and the PVN (a primary site for the coordination of central catabolic signals)<sup>20</sup>, is consistent with a role for OEA as a peripheral regulator of feeding behaviour.

The anorexic effects of OEA are reminiscent of those produced by gut peptides such as CCK. Because the release of CCK from duodenum is regulated by nutrients<sup>22</sup>, we studied the impact of feeding on intestinal OEA biosynthesis. Analyses with high-performance liquid chromatography and mass spectrometry (HPLC/MS) revealed that small-intestinal tissue from free-feeding rats contains substantial amounts of OEA (354 ± 86 pmol g<sup>-1</sup>,  $n = 3$ ). Intestinal OEA levels were markedly decreased after food deprivation, but returned to baseline after refeeding (Fig. 5a). By contrast, no such changes were observed in stomach (control, 210 ± 20 pmol g<sup>-1</sup>; starvation, 238 ± 84 pmol g<sup>-1</sup>; starvation–refeeding, 239 ± 60 pmol g<sup>-1</sup>;  $n = 3$ ). Variations in intestinal OEA levels were accompanied by parallel alterations in NAT activity, which participates in OEA formation<sup>4</sup>, but not in FAAH activity, which catalyses OEA hydrolysis (Fig. 5b, c)<sup>11,12</sup>. These findings suggest that starvation and feeding reciprocally regulate OEA biosynthesis in small intestine. In agreement with an intra-abdominal source of OEA, we found that plasma OEA levels in starved rats are higher in portal than in caval blood (porta, 14.6 ± 1.8 pmol ml<sup>-1</sup>; cava, 10.3 ± 2.8 pmol ml<sup>-1</sup>;  $n = 5$ ). The contribution of other intra-abdominal tissues to OEA formation cannot be excluded at present.

Our results suggest a hypothetical model for the role of OEA in feeding behaviour. According to this model, food intake may stimulate NAT activity, enhancing OEA biosynthesis in the small intestine and possibly other intra-abdominal tissues. Newly produced OEA may activate local sensory fibres, which may in turn inhibit feeding by engaging brain structures such as the NST and PVN. A number of questions remain, such as what physiological stimuli initiate and terminate OEA biosynthesis, whether there is a functional relationship between OEA and other nutritional signals, and what the molecular targets of OEA are. Irrespective of the answers, our results reveal an unexpected role for OEA in the peripheral regulation of feeding, and provide a framework to develop medicines for the treatment of eating disorders. □

## Methods

### Animals

We used male Wistar rats (200–350 g). All procedures met the National Institutes of Health guidelines for the care and use of laboratory animals, and the European Communities directive 86/609/EEC regulating animal research.

### Chemicals

FAEs and [<sup>3</sup>H] FAEs were synthesized in the laboratory<sup>23</sup>; 1,2-dioleoyl-*sn*-glycero-phosphoethanolamine-*N*-oleyl was purchased from Avanti Polar Lipids; SR141716A was provided by RBI as part of the Chemical Synthesis Program of the National Institute of Mental Health; SR144528 was a gift of Sanofi Recherche; all other drugs were from Tocris or Sigma. FAEs were dissolved in dimethylsulphoxide (DMSO) and administered in 70% DMSO in sterile saline (acute treatments) or 5% Tween 80 with 5% propylenglycol in sterile saline (subchronic treatments) (1 ml kg<sup>-1</sup> i.p.). Capsaicin was administered in 10% Tween 80, 10% ethanol and 80% saline; SR141716A, SR144528, CCK-8 and CP-93129 in 5% Tween 80, 5% propylenglycol and 90% saline (1 ml kg<sup>-1</sup> i.p.).

### Enzyme assays

In all biochemical experiments, rats were killed and tissues collected between 14:00 and 16:00. Microsome fractions were prepared as described<sup>24</sup>. NAT assays were performed using 1,2-di[<sup>14</sup>C]palmityl-*sn*-glycerophosphocholine as a substrate (108 mCi mmol<sup>-1</sup>, Amersham) under conditions that were linear with time and protein concentrations<sup>5</sup>. FAAH assays were performed under linear conditions according to ref. 24, except that [<sup>3</sup>H]anandamide (arachidonyl-[1-<sup>3</sup>H]ethanolamide; 60 Ci mmol<sup>-1</sup>; ARC) was included as a substrate and radioactivity was measured in the aqueous phase after chloroform extraction.

### HPLC/MS analyses

FAEs were extracted from tissues with a methanol–chloroform mixture and fractionated by column chromatography<sup>23</sup>. FAEs were quantified by HPLC/MS, with an isotope dilution method<sup>25</sup>.

### Blood chemistry

We used commercial kits to measure glucose, insulin, leptin and other plasma metabolites and enzymes (Sigma and Linco Research). Plasma prolactin, corticosterone and luteinizing hormone were quantified by radioimmunoassay<sup>26</sup>.

### Feeding experiments

For acute experiments, we measured food intake in rats deprived of food for 24 h (ref. 27) that were habituated to the experimental setting. We administered drugs 15 min before food presentation. For subchronic experiments, freely fed rats received vehicle injections for two days. On day 3, we divided the animals into two equal groups and gave them daily injections of vehicle or OEA (5 mg kg<sup>-1</sup> at 19:00) for seven consecutive days, while measuring body mass, food intake and water intake. A third group of rats (pair fed) received an amount of food identical to that consumed by the OEA group. On day 9 of the study, the animals were killed and plasma was collected for biochemical analyses.

### Conditioned taste aversion

Rats were deprived of water for 24 h and then accustomed to drinking from a graded bottle during a 30-min test period for 4 days. On day 5, water was substituted with a 0.1% saccharin solution and, 30 min later, the animals received injections of vehicle, OEA (20 mg kg<sup>-1</sup>) or lithium chloride (0.4 M, 7.5 ml kg<sup>-1</sup>). During the following 2 days, water consumption was recorded over 30-min test periods. The animals were then presented with water or saccharin, and drinking was measured.

### Operant responses for food

Rats were trained to press a lever for food on a fixed ratio (FR) 1 schedule of reinforcement, while restricted to 20 g of food per rat per day. Once a stable response was achieved, the animals were trained to acquire an FR5, time out 2-min schedule of food reinforcement and kept with limited access to food. When a stable baseline was obtained, the animals were used to test the effects of vehicle or OEA (5 or 20 mg kg<sup>-1</sup>) administered 15 min before lever presentation. Tests lasted 60 min.

### Other behavioural assays

The elevated plus maze test was conducted as described<sup>26</sup> after the administration of vehicle or OEA (20 mg kg<sup>-1</sup> i.p.). Horizontal activity in an open field and pain threshold in the hot plate test (55 °C) were measured 15 min after injection of vehicle or OEA (20 mg kg<sup>-1</sup>). Rectal temperature was measured with a digital thermometer.

### In situ hybridization

We accustomed rats to the handling and injection procedure for 5 days. On day 6, we administered vehicle, OEA (10 mg kg<sup>-1</sup> i.p.), or oleic acid (10 mg kg<sup>-1</sup>) and killed the rats 60 min later by decapitation under anaesthesia. *In situ* hybridization analyses were conducted using [<sup>35</sup>S]-labelled antisense RNA probes for *c-fos*<sup>28</sup> and choline acetyl transferase<sup>29</sup>. Average hybridization densities were determined from at least three tissue sections per rat. Statistical significance was evaluated using a one-way analysis of variance (ANOVA) followed by the Tukey–Kramer post-hoc test for paired comparisons.

## Data analysis

Results are expressed as mean  $\pm$  s.e.m. of  $n$  separate experiments. The significance of differences among groups was evaluated using ANOVA followed by a Student–Newman–Keuls *post-hoc* test, unless indicated otherwise.

Received 5 June; accepted 12 September 2001.

- Di Marzo, V. *et al.* Formation and inactivation of endogenous cannabinoid anandamide in central neurons. *Nature* **372**, 686–691 (1994).
- Piomelli, D., Beltramo, M., Giuffrida, A. & Stella, N. Endogenous cannabinoid signaling. *Neurobiol. Dis.* **5**, 462–473 (1998).
- Bachur, N. R., Masek, K., Melmon, K. L. & Udenfriend, S. Fatty acid amides of ethanolamine in mammalian tissues. *J. Biol. Chem.* **240**, 1019–1024 (1965).
- Schmid, H. H., Schmid, P. C. & Natarajan, V. The *N*-acylation-phosphodiesterase pathway and cell signalling. *Chem. Phys. Lipids* **80**, 133–142 (1996).
- Chapman, K. D. Emerging physiological roles for *N*-acylphosphatidylethanolamine metabolism in plants: signal transduction and membrane protection. *Chem. Phys. Lipids* **108**, 221–229 (2000).
- Giuffrida, A. *et al.* Dopamine activation of endogenous cannabinoid signaling in dorsal striatum. *Nature Neurosci.* **2**, 358–363 (1999).
- Berdyshev, E. V., Schmid, P. C., Dong, Z. & Schmid, H. H. Stress-induced generation of *N*-acylethanolamines in mouse epidermal JB6 P+ cells. *Biochem. J.* **346**, 369–374 (2000).
- Cadas, H., Gailllet, S., Beltramo, M., Venance, L. & Piomelli, D. Biosynthesis of an endogenous cannabinoid precursor in neurons and its control by calcium and cAMP. *J. Neurosci.* **16**, 3934–3942 (1996).
- Cadas, H., di Tomaso, E. & Piomelli, D. Occurrence and biosynthesis of endogenous cannabinoid precursor, *N*-arachidonoyl phosphatidylethanolamine, in rat brain. *J. Neurosci.* **17**, 1226–1242 (1997).
- Beltramo, M. *et al.* Functional role of high-affinity anandamide transport, as revealed by selective inhibition. *Science* **277**, 1094–1097 (1997).
- Schmid, P. C., Zuzarte-Augustin, M. L. & Schmid, H. H. Properties of rat liver *N*-acylethanolamine amidohydrolase. *J. Biol. Chem.* **260**, 14145–14149 (1985).
- Cravatt, B. F. *et al.* Molecular characterization of an enzyme that degrades neuromodulatory fatty-acid amides. *Nature* **384**, 83–87 (1996).
- Devane, W. A. *et al.* Isolation and structure of a brain constituent that binds to the cannabinoid receptor. *Science* **258**, 1946–1949 (1992).
- Williams, C. M. & Kirkham, T. C. Anandamide induces overeating: mediation by central cannabinoid (CB1) receptors. *Psychopharmacology* **143**, 315–317 (1999).
- Khanolkar, A. D. & Makriyannis, A. Structure–activity relationships of anandamide, an endogenous cannabinoid ligand. *Life Sci.* **65**, 607–616 (1999).
- Pertwee, R. G. Cannabinoid receptor ligands: clinical and neuropharmacological considerations, relevant to future drug discovery and development. *Exp. Opin. Invest. Drugs* **9**, 1553–1571 (2000).
- Kaneko, H., Kaunitz, J. & Tache, Y. Vagal mechanisms underlying gastric protection induced by chemical activation of raphe pallidus in rats. *Am. J. Physiol.* **275**, G1056–G1062 (1998).
- MacLean, D. B. Abrogation of peripheral cholecystokinin-satiety in the capsaicin treated rat. *Regul. Pept.* **11**, 321–333 (1985).
- Lee, M. D., Aloyo, V. J., Fluharty, S. J. & Simansky, K. J. Infusion of the serotonin1B (5-HT1B) agonist CP-93,129 into the parabrachial nucleus potently and selectively reduces food intake in rats. *Psychopharmacology* **136**, 304–307 (1998).
- Schwartz, M. W., Woods, S. C., Porte, D. J., Seeley, R. J. & Baskin, D. G. Central nervous system control of food intake. *Nature* **404**, 661–671 (2000).
- Curran, T., Gordon, M. B., Rubino, K. L. & Sambucetti, L. C. Isolation and characterization of the *c-fos* (rat) cDNA and analysis of post-translational modification *in vitro*. *Oncogene* **2**, 79–84 (1987).
- Ritter, R. C., Covasa, M. & Matson, C. A. Cholecystokinin: proofs and prospects for involvement in control of food intake and body weight. *Neuropeptides* **33**, 387–399 (1999).
- Giuffrida, A. & Piomelli, D. In *Lipid Second Messengers* (eds Laychock, S. G. & Rubin, R. P.) 113–133 (CRC, Boca Raton, 1998).
- Désarnaud, F., Cadas, H. & Piomelli, D. Anandamide amidohydrolase activity in rat brain microsomes. Identification and partial characterization. *J. Biol. Chem.* **270**, 6030–6035 (1995).
- Giuffrida, A., Rodríguez de Fonseca, F. & Piomelli, D. Quantification of bioactive acylethanolamides in rat plasma by electrospray mass spectrometry. *Anal. Biochem.* **280**, 87–93 (2000).
- Navarro, M. *et al.* Acute administration of the CB1 cannabinoid receptor antagonist SR 141716A induces anxiety-like responses in the rat. *Neuroreport* **8**, 491–496 (1997).
- Navarro, M. *et al.* Colocalization of glucagon-like peptide-1 (GLP-1) receptors, glucose transporter GLUT-2, and glucokinase mRNAs in rat hypothalamic cells: evidence for a role of GLP-1 receptor agonists as an inhibitory signal for food and water intake. *J. Neurochem.* **67**, 1982–1991 (1996).
- Guthrie, K. M., Anderson, A. J., Leon, M. & Gall, C. Odor-induced increases in *c-fos* mRNA expression reveal an anatomical “unit” for odor processing in olfactory bulb. *Proc. Natl Acad. Sci. USA* **90**, 3329–3333 (1993).
- Lauterborn, J. C., Isackson, P. J., Montalvo, R. & Gall, C. M. *In situ* hybridization localization of choline acetyltransferase mRNA in adult rat brain and spinal cord. *Brain Res. Mol. Brain Res.* **17**, 59–69 (1993).

**Supplementary Information** accompanies the paper on Nature’s website (<http://www.nature.com>).

## Acknowledgements

We thank M. Schwartz for comments; R. Carrera and A. M. Basso for help with initial experiments; T. Dinh, M. Guzmán and C. Sánchez for reading the manuscript critically; M. A. Gorriti, J. Muñoz, F. Désarnaud, M. A. Villanúa and Y. Xie for help with experiments; and F. Valiño for editorial assistance. This research was supported by grants from the National Institute of Drug Abuse (to D.P.) and from the National Institute of Mental Health (to C.G.). Additional support was from the Comunidad de Madrid, Del Amo Program and Plan Nacional sobre Drogas (to F.R.F. and M.N.). F.R.F. is a research fellow of the Jaime del Amo Program, Complutense University.

Correspondence and requests for materials should be addressed to D.P. (e-mail: piomelli@uci.edu).

# A subset of NSAIDs lower amyloidogenic A $\beta$ 42 independently of cyclooxygenase activity

Sascha Weggen\*, Jason L. Eriksen†, Pritam Das†, Sarah A. Sagi\*, Rong Wang‡, Claus U. Pietrzik\*, Kirk A. Findlay†, Tawnya E. Smith†, Michael P. Murphy†, Thomas Bulter§, David E. Kang\*, Numa Marquez-Sterling||, Todd E. Golde† & Edward H. Koo\*

\* Department of Neurosciences, University of California San Diego, La Jolla, California 92093, USA

† Department of Neuroscience and Pharmacology, Mayo Clinic Jacksonville, Jacksonville, Florida 32224, USA

‡ Department of Human Genetics, Mount Sinai School of Medicine, New York, New York 10029, USA

§ Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, USA

|| Department of Pathology, Northwestern University Medical School, Chicago, Illinois 60611, USA

Epidemiological studies have documented a reduced prevalence of Alzheimer’s disease among users of nonsteroidal anti-inflammatory drugs (NSAIDs)<sup>1–5</sup>. It has been proposed that NSAIDs exert their beneficial effects in part by reducing neurotoxic inflammatory responses in the brain, although this mechanism has not been proved. Here we report that the NSAIDs ibuprofen, indomethacin and sulindac sulphide preferentially decrease the highly amyloidogenic A $\beta$ 42 peptide (the 42-residue isoform of the amyloid- $\beta$  peptide) produced from a variety of cultured cells by as much as 80%. This effect was not seen in all NSAIDs and seems not to be mediated by inhibition of cyclooxygenase (COX) activity, the principal pharmacological target of NSAIDs<sup>6</sup>. Furthermore, short-term administration of ibuprofen to mice that produce mutant  $\beta$ -amyloid precursor protein (APP) lowered their brain levels of A $\beta$ 42. In cultured cells, the decrease in A $\beta$ 42 secretion was accompanied by an increase in the A $\beta$ (1–38) isoform, indicating that NSAIDs subtly alter  $\gamma$ -secretase activity without significantly perturbing other APP processing pathways or Notch cleavage. Our findings suggest that NSAIDs directly affect amyloid pathology in the brain by reducing A $\beta$ 42 peptide levels independently of COX activity and that this A $\beta$ 42-lowering activity could be optimized to selectively target the pathogenic A $\beta$ 42 species.

Increasing evidence suggests that a key event in the pathogenesis of Alzheimer’s disease is the altered production, aggregation and deposition of the A $\beta$  peptide, a proteolytic fragment of 40–42 residues derived from APP. The longer isoform, A $\beta$ 42, is selectively increased in all presenilin mutations analysed and in most APP mutations that cause early-onset familial Alzheimer’s disease. A $\beta$ 42 is the A $\beta$  species initially deposited in brain, and is particularly prone to aggregation *in vitro*. Therefore, A $\beta$ 42 is believed by many to be the main culprit in the pathogenesis of Alzheimer’s disease<sup>7</sup>.

We examined whether NSAIDs alter APP processing and generation of A $\beta$ , particularly the A $\beta$ 42 species. We treated cells with increasing concentrations of various NSAIDs, and analysed A $\beta$ 40 and A $\beta$ 42 levels in culture medium by sensitive sandwich enzyme-link immunosorbent assay (ELISA) as described previously<sup>8</sup>. The range of NSAID concentrations was chosen on the basis of tolerated plasma concentrations achieved in humans<sup>9</sup>. Multiple NSAIDs were examined in this study, several of them sharing similar activities (see below; an overview describing the cell lines and compounds studied is provided in the Supplementary Information). For brevity, the non-selective COX-inhibitor sulindac sulphide, the active metabolite of the pro-drug sulindac with well recognized antineoplastic

properties<sup>10</sup>, is presented as the prototypic NSAID in its striking preferential inhibition of A $\beta$ 42 secretion. In Chinese hamster ovary (CHO) cells stably transfected with both APP751 (wild-type APP, WT-APP) and the PS1 mutant M146L (PS1-M146L), a 50–70% reduction in the A $\beta$ 42/A $\beta$ 40 quotient was achieved at concentrations of 40–60  $\mu$ M without significant reduction in total A $\beta$  (A $\beta$ 40 + A $\beta$ 42) levels (Fig. 1a). This result was confirmed in CHO cells overexpressing WT-APP and in CHO cells transfected with the APP V717F mutation (see Supplementary Information). To exclude the possibility that this effect is specific to a particular cell type, we examined A $\beta$  secretion in response to sulindac sulphide treatment in the human neuroglioma cell line HS683 stably transfected with APP695. A dose-dependent inhibition of A $\beta$ 42 secretion similar to that seen with CHO cells was observed (Fig. 1b). Sulindac sulphide also reduced A $\beta$ 42 secretion in kidney HEK293 cells, H4 neuroglioma cells, and primary mouse embryonic fibroblasts (Supplementary Information). Thus, this effect was observed in cell lines of rodent and human origin and is not dependent on cell type.

Ibuprofen reduced amyloid plaque pathology in a mouse model of Alzheimer's disease, and indomethacin seemed to slow the cognitive decline in patients with Alzheimer's disease in a placebo-controlled pilot study<sup>3,11</sup>. These two non-selective COX inhibitors also decreased the A $\beta$ 42/A $\beta$ 40 quotient by selective inhibition of A $\beta$ 42 secretion in a dose-dependent manner. In WT-APP PS1-M146L CHO cells, a 50% reduction in the A $\beta$ 42/A $\beta$ 40 quotient was reached at ibuprofen concentrations of 200–300  $\mu$ M (Fig. 1c) and at indomethacin concentrations of 25–50  $\mu$ M (Fig. 1d). Similarly, total A $\beta$  levels were not significantly affected with either ibuprofen (up to 500  $\mu$ M) or indomethacin (up to

**Table 1 Acute dosing with ibuprofen lowers brain A $\beta$ 42**

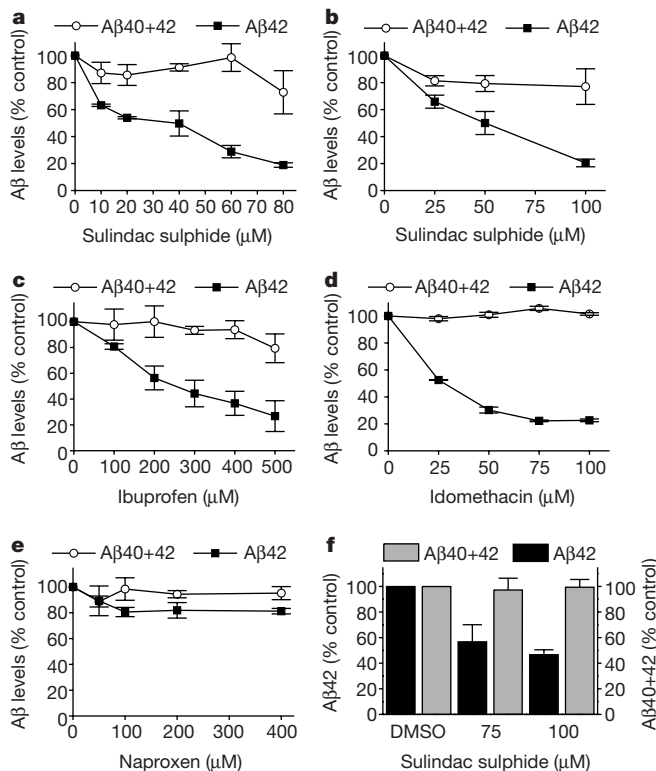
| Treatment | <i>n</i> | A $\beta$ 40 (pmol g <sup>-1</sup> ) | A $\beta$ 42 (pmol g <sup>-1</sup> ) | A $\beta$ 42 (% of A $\beta$ 40 + A $\beta$ 42) |
|-----------|----------|--------------------------------------|--------------------------------------|-------------------------------------------------|
| Control   | 18       | 18.5 $\pm$ 2.8                       | 7.7 $\pm$ 1.3                        | 29.5 $\pm$ 3.4                                  |
| Naproxen  | 7        | 18.6 $\pm$ 1.1                       | 7.9 $\pm$ 0.6                        | 29.8 $\pm$ 1.6                                  |
| Ibuprofen | 15       | 17.5 $\pm$ 1.7                       | 4.7 $\pm$ 2.0*                       | 20.8 $\pm$ 7.5*                                 |

Tg2576 mice were dosed as described in the text. Values are expressed as mean  $\pm$  s.d.  
\**P* < 0.01 by Dunnett's test.

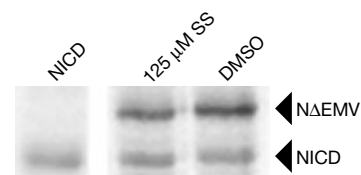
150  $\mu$ M) (Fig. 1c and d). No toxicity was detected by standard MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide) assay in CHO or HS683 cells or by [<sup>3</sup>H]-thymidine incorporation for CHO cells treated with sulindac sulphide concentrations up to 100  $\mu$ M, with indomethacin up to 200  $\mu$ M, or with ibuprofen up to 1 mM (data not shown). Therefore, the A $\beta$ 42 effect is not related to cytotoxicity.

We next examined sulindac sulphone, a second oxidative metabolite of sulindac with potent antineoplastic effects that lacks COX-inhibitory activity<sup>12</sup>. Unlike sulindac sulphide, sulindac sulphone had no effect on the A $\beta$ 42/A $\beta$ 40 quotient in WT-APP CHO cells with concentrations up to 400  $\mu$ M (data not shown). This finding suggests a role for COX inhibition in the preferential reduction of A $\beta$ 42 secretion by NSAIDs. We therefore investigated whether reduction of A $\beta$ 42 levels is a characteristic common to all NSAIDs by assessing several other clinically approved NSAIDs. Naproxen is a non-selective COX inhibitor that belongs to the same structural class as ibuprofen. However, unlike ibuprofen or sulindac sulphide, treatment of WT-APP CHO cells with naproxen at concentrations up to 400  $\mu$ M did not lower either the A $\beta$ 42/A $\beta$ 40 quotient or total A $\beta$  levels (Fig. 1e). Similar negative results were obtained with aspirin, meloxicam (an NSAID with preferential COX-2 activity), SC-560 (a selective COX-1 inhibitor) and celecoxib (a selective COX-2 inhibitor) (see Supplementary Information). These data demonstrated that the capacity to lower A $\beta$ 42 secretion is not associated with all NSAIDs. Furthermore, the effective NSAID concentrations used in our experiments were clearly in excess of the levels required for complete inhibition of COX-1 and COX-2 in cell-based assays<sup>13</sup> (see Supplementary Information), indicating that the reduction in A $\beta$ 42 levels may be independent of COX activity.

To demonstrate that the decrease in A $\beta$ 42 levels is not mediated by inhibition of COX activity and concomitant reduction in prostaglandin synthesis, fibroblasts deficient in both COX-1 and COX-2 by targeted gene deletions (COX-1<sup>-/-</sup> COX-2<sup>-/-</sup>)<sup>14</sup> (see Supplementary Information) were examined for the effect of NSAIDs on A $\beta$ 42 levels. The basal levels of A $\beta$  peptides were unchanged compared with littermate control mouse fibroblasts, indicating that the loss of COX-1 and COX-2 enzymes did not by itself alter A $\beta$ 42 levels (data not shown). However, treatment with sulindac sulphide reduced A $\beta$ 42 levels and altered the A $\beta$ 42/A $\beta$ 40 quotient in a similar fashion to that seen in CHO and HS683 neuroglioma cells (Fig. 1f) and identical to littermate control



**Figure 1** Analysis of A $\beta$  from cultured cells treated with NSAIDs by ELISA. A $\beta$ 42 and total A $\beta$  levels (A $\beta$ 40 plus A $\beta$ 42 values, A $\beta$ 40+42) were normalized to values obtained from vehicle-treated cells. Treatment with sulindac sulphide preferentially reduced A $\beta$ 42 levels in the medium of WT-APP PS1-M146L CHO cells (a) and human neuroglioma cells (b) in a dose-dependent manner. Selective inhibition of A $\beta$ 42 levels was also observed with ibuprofen (c) and indomethacin (d), but not with naproxen (e) in WT-APP CHO cells. Sulindac sulphide similarly reduced A $\beta$ 42 levels in fibroblasts deficient in COX-1 and COX-2 (f). Each panel shows the mean  $\pm$  s.d. of all experiments.



**Figure 2** Analysis of Notch processing after treatment with sulindac sulphide in HEK 293 cells. NICD formation from a constitutively cleaved Notch construct (N $\Delta$ EMV) was not impaired in cells treated with sulindac sulphide (125  $\mu$ M SS). The left lane shows control transfection with an NICD construct to identify the cleavage fragment.

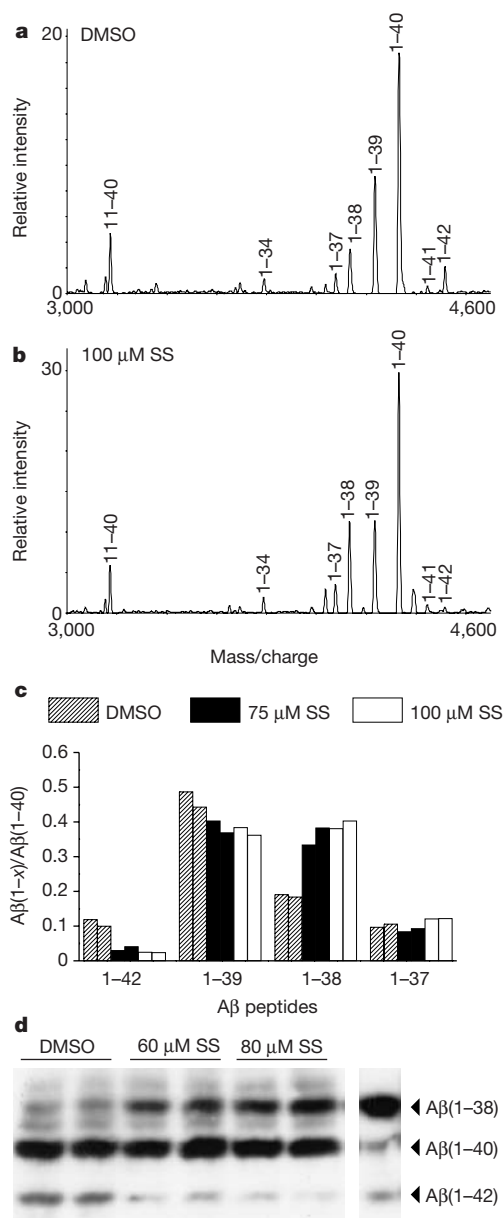


fibroblasts (data not shown). These results provide compelling evidence that the reduction in A $\beta$ 42 is not mediated by inhibition of COX activity, the principal mode of action of NSAIDs. Whether other NSAID-responsive pathways mediate this activity, such as by modulating lipoxygenase activity or transcription controlled by peroxisome proliferator-activated receptors (PPARs), is unknown.

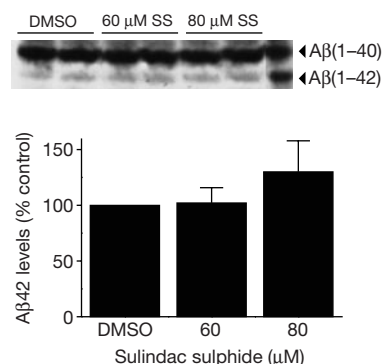
In a recent study, chronic high-dose ibuprofen treatment was shown to significantly reduce amyloid pathology, neuritic dystrophy, plaque-associated gliosis and IL-1 expression in Tg2576 transgenic mice<sup>11</sup>. After 6 months of treatment, amyloid plaque numbers

and A $\beta$  levels in brain were reduced almost 50% and 40%, respectively. To determine whether the COX-independent reduction in A $\beta$ 42 seen with some NSAIDs in cultured cells could in part account for the long-term effects that have been reported, we examined whether treatment of Tg2576 mice with ibuprofen preferentially reduces A $\beta$ 42 levels. For these studies, 3-month-old female Tg2576 mice were orally dosed with 50 mg kg<sup>-1</sup> d<sup>-1</sup> of ibuprofen ( $n = 15$ ) or 50 mg kg<sup>-1</sup> d<sup>-1</sup> of naproxen ( $n = 7$ ), or mock treated ( $n = 18$ ), and brain levels of SDS-soluble A $\beta$ 40 and A $\beta$ 42 were measured by ELISA. Administration of ibuprofen resulted in a highly significant 39% decrease in levels of A $\beta$ 42 without any changes in A $\beta$ 40 compared with mock-treated animals (Table 1 and Supplementary Information). Consistent with the tissue culture results, treatment with naproxen did not alter the levels of A $\beta$ 42 in brain at a dose that has been shown to inhibit prostaglandin production in rodent brain<sup>15</sup>. These results strongly indicate the physiological relevance of our *in vitro* studies. In the previous study<sup>11</sup>, a reduction in parenchymal but not vascular amyloid deposits was noted. Because vascular A $\beta$  deposits consist primarily of A $\beta$ 40, whereas parenchymal deposits are a mixture of A $\beta$ 40 and A $\beta$ 42, the lack of effect on vascular amyloid is in agreement with our hypothesis that ibuprofen treatment may prevent amyloid pathology by decreasing the A $\beta$ 42/A $\beta$ 40 quotient in brain.

NSAIDs are unique in their ability to change the A $\beta$ 42/A $\beta$ 40 quotient by selectively decreasing A $\beta$ 42 levels. Recently,  $\gamma$ -secretase inhibitors have been developed that either inhibit the generation of all A $\beta$  species or show preferential activity against A $\beta$ 40. Several of these inhibitors have been demonstrated to bind presenilins, leading to the hypothesis that presenilins are actual  $\gamma$ -secretases<sup>16</sup>. Because presenilins are essential not only for  $\gamma$ -secretase cleavage of APP but also for proteolytic processing of the Notch receptor<sup>17</sup>,  $\gamma$ -secretase inhibitors may concomitantly inhibit Notch cleavage, resulting in adverse side effects<sup>18</sup>. Furthermore, treatment with  $\gamma$ -secretase inhibitors leads to the accumulation of high amounts of carboxy-terminal fragments (CTFs) of APP that may be neurotoxic<sup>19,20</sup>. We therefore examined multiple parameters in APP processing as well as Notch intramembrane cleavage to determine whether any of these cellular pathways are altered in response to NSAID treatment. WT-APP CHO cells were treated with increasing concentrations of sulindac sulphide as described above. Western blot analysis with an APP C-terminal polyclonal antibody did not show any changes in the levels of full-length APP or APP CTF species cleaved by  $\alpha$ - or  $\beta$ -secretase (see Supplementary Information). Similarly, secretion of



**Figure 3** A $\beta$  species in medium after NSAID treatment. **a**, Representative mass spectra of A $\beta$  peptides from WT-APP CHO cells treated with 100  $\mu$ M sulindac sulphide (SS) or DMSO vehicle showing a decrease in A $\beta$ (1-42) species but an increase in A $\beta$ (1-38) species. **b**, Quantitative mass spectrometry analysis of A $\beta$  peptide levels calculated as a quotient, A $\beta$ (1-x)/A $\beta$ (1-40). **c**, The changes in A $\beta$ (1-42) and A $\beta$ (1-38) levels in medium were confirmed by bicine/urea SDS-PAGE. Standard A $\beta$ (1-38), A $\beta$ (1-40) and A $\beta$ (1-42) peptides were used to identify A $\beta$  species (right lane). The lane with control peptides was intentionally offset because the exposure time was longer to provide comparable band intensities. **d**, Reduction in A $\beta$ (1-42)/A $\beta$ (1-40) was accompanied by an increase in A $\beta$ (1-38)/A $\beta$ (1-40). Duplicate measurements are shown.



**Figure 4** Turnover of A $\beta$ 42 after NSAID treatment. A $\beta$ 40 and A $\beta$ 42 were assayed from untransfected CHO cells incubated with conditioned medium obtained from CHO cells transfected with WT-APP and treated with increasing concentrations of sulindac sulphide (SS) (see Methods). The levels of A $\beta$ 40 and A $\beta$ 42 from the conditioned medium were virtually identical to vehicle-treated controls analysed by western blotting (top) or ELISA (bottom), indicating that A $\beta$  turnover (clearance or degradation) was not affected by NSAID treatment. The ELISA results represent mean  $\pm$  s.d. of all experiments.

the APP ectodomain (APPs) was not altered, as determined by western blotting analysis of conditioned medium using two different APP monoclonal antibodies (see Supplementary Information). Furthermore, by pulse-chase metabolic labelling, APP turnover was unchanged after treatment with sulindac sulphide (see Supplementary Information).

We next examined APP internalization because processing of APP in the endocytic pathway is hypothesized to be a major route for the generation and secretion of both A $\beta$ 40 and A $\beta$ 42 (ref. 21). Using a previously described APP internalization assay<sup>21</sup>, we found the ratio of cell surface APP to internalized APP to be virtually identical in cells treated with sulindac sulphide compared with vehicle-treated cells (see Supplementary Information). Finally, we analysed Notch intramembrane cleavage and formation of the Notch intracellular cytoplasmic domain (NICD) in HEK293 cells after transfection with the Myc-tagged N $\Delta$ EMV complementary DNA, an artificial Notch receptor construct that undergoes constitutive cleavage<sup>22</sup>. Consistent with the results for APP CTFs shown above, treatment with sulindac sulphide did not impair Notch cleavage or NICD formation (Fig. 2). Similarly, treatment with ibuprofen (500  $\mu$ M) or indomethacin (150  $\mu$ M) did not result in either accumulation of APP CTFs or inhibition of Notch cleavage (data not shown). These results demonstrate that NSAID treatment did not significantly perturb APP processing or  $\gamma$ -secretase activity. However, the preferential decrease in A $\beta$ 42 could be explained by minor changes in  $\gamma$ -secretase activity that were not detectable in the preceding assays.

To determine how NSAIDs may alter A $\beta$ 42 levels, two additional sets of experiments were performed. First, we analysed the full spectrum of A $\beta$  species secreted by cells treated with sulindac sulphide by immunoprecipitation with mass spectrometry (IPMS). In contrast to ELISA measurements, IPMS provides definitive information about the length and identity of the individual A $\beta$  peptides, for example A $\beta$ (1–38), A $\beta$ (1–40), and so on. The IPMS results extended our findings in significant ways. As expected, treatment with 75–100  $\mu$ M sulindac sulphide reduced the level of A $\beta$ (1–42) whereas A $\beta$ (1–40) was essentially unaffected. Remarkably, the decrease in A $\beta$ (1–42) secretion was accompanied by a dose-dependent increase in A $\beta$ (1–38) species, resulting in a two-fold increase in the A $\beta$ (1–38)/A $\beta$ (1–40) quotient (Fig. 3a and b). Other A $\beta$  peptide species (A $\beta$ (1–37) and A $\beta$ (1–39)) did not show consistent changes after treatment. To confirm the IPMS results, A $\beta$  was immunoprecipitated from conditioned media of WT-APP PS1-M146L CHO cells and fractionated on a gel system that allows the resolution of individual A $\beta$  species<sup>23</sup>. As anticipated, a large reduction in an immunoreactive band corresponding to A $\beta$ (1–42) was accompanied by a dose-dependent increase in an immunoreactive species corresponding to A $\beta$ (1–38) (Fig. 3c).

Two potential mechanisms may explain this unprecedented change in A $\beta$  production after NSAID treatment. Sulindac sulphide and other NSAIDs could reduce A $\beta$ 42 secretion by shifting  $\gamma$ -secretase activity towards production of A $\beta$ 38. Alternatively, NSAIDs may stimulate the activity of an exopeptidase that converts A $\beta$ 42 into shorter A $\beta$  species such as A $\beta$ 38. In the second experiment, we therefore asked whether the turnover of A $\beta$ 42 in culture medium could be enhanced after NSAID treatment. Up to this point, only the levels of A $\beta$  peptides had been assessed and it is possible that NSAIDs may facilitate the degradation or cell-mediated clearance of A $\beta$ 42 (refs 24–26). To examine this possibility, untransfected CHO cells were cultured in conditioned medium obtained from WT-APP CHO cells supplemented with increasing concentrations of sulindac sulphide. Because untransfected CHO cells do not produce A $\beta$  to any significant level, all the measurable A $\beta$  peptides from this culture paradigm were derived exogenously. Under this condition, sulindac sulphide did not reduce the levels of A $\beta$ 42 compared with controls (Fig. 4), indicating that NSAIDs did not selectively target A $\beta$ 42 for catabolism, either by activating

exopeptidase activity or by selective clearance. Future studies will be required to address this complex mechanism as there are no precedents for this pharmacological alteration in A $\beta$  C-terminal cleavages, putatively attributable to altered  $\gamma$ -secretase activity.

Our observations suggest a mechanism through which NSAIDs might confer protection from Alzheimer's disease that is independent of direct anti-inflammatory properties. Because not all NSAIDs lower A $\beta$ 42 levels, this A $\beta$ 42 activity may be an important criterion to consider in clinical trials of NSAIDs for the treatment of Alzheimer's disease. Our results suggest an explanation for the variable data obtained from epidemiological studies and recent negative clinical results with COX-2-selective NSAIDs<sup>27</sup>, as we find that many NSAIDs (including these selective COX-2 inhibitors) lack the ability to lower A $\beta$ 42. However, our results do not exclude the potential benefit of NSAIDs in reducing the inflammatory response in the Alzheimer's disease brain. Finally, in contrast to the current generation of  $\gamma$ -secretase inhibitors, NSAIDs do not perturb APP or Notch processing but rather seem to induce a subtle shift in  $\gamma$ -secretase activity. However, significant gastrointestinal and renal toxicity associated with long-term COX-1 inhibition limit the clinical utility of current NSAIDs as A $\beta$ 42-lowering agents. Because the A $\beta$ 42 effect we described is independent of COX activity, compounds with optimized A $\beta$ 42 reduction and little to no inhibition of COX-1 activity are likely to be identified. Such agents would represent a new generation of 'anti-amyloid' drugs that selectively target production of the highly amyloidogenic A $\beta$ 42 species without inhibiting either COX activity or the vital physiological functions of  $\gamma$ -secretase. □

## Methods

### Cell culture and drug treatment

CHO cells stably transfected with human APP751 or transfected with both human APP751 and human mutant PS1 (M146L), CHO cells transfected with human mutant APP751 (V717F), human neuroglioma cells HS683 transfected with APP695, HEK 293 cells transfected with APP695, and spontaneously immortalized embryonic fibroblasts from COX1- and COX-2-deficient animals were maintained in DMEM medium supplemented with 10% fetal bovine serum and 100 U ml<sup>-1</sup> penicillin/streptomycin (Life Technologies). The NSAIDs sulindac sulphide (50 mM, Biomol), sulindac sulphone (50 mM, Biomol), naproxen (100 mM, Cayman Chemical), aspirin (2.5 M, ICN Bio-medicals), meloxicam (50 mM, Calbiochem) and SC-560 (50 mM, Calbiochem) were dissolved in dimethyl sulphoxide (DMSO). Indomethacin (50 mM, Biomol) and (S)-ibuprofen (250 mM, Biomol) were dissolved in ethanol. Celecoxib was prepared from commercial Celebrex capsules (Searle) by solvent extraction followed by recrystallization. Nuclear magnetic resonance (NMR) and mass spectrometry verified the identity and purity of celecoxib. Purified celecoxib was dissolved in ethyl acetate (10 mM). For analysis of A $\beta$  secretion, APP processing and Notch cleavage, cells were cultured in serum-containing medium and pretreated overnight with the specific NSAIDs. Medium was changed and treatment was continued for another 24 h. All experiments were repeated 2–4 times in duplicate or triplicate and results either from a representative experiment or from all experiments (mean  $\pm$  s.d.) are shown.

### ELISA

A $\beta$  was analysed by sandwich ELISA as described previously<sup>8</sup>. Media were collected after conditioning for 24 h, and cell debris was removed by centrifugation. Complete protease inhibitor cocktail (Roche) was added and A $\beta$ 40 and A $\beta$ 42 were quantified by BAN50/BA27 and BAN50/BC05 ELISAs or 3160/BA27 and 3160/BC05 ELISAs. All measurements were performed in duplicate.

### Adenoviral infection of COX1<sup>-/-</sup> COX-2<sup>-/-</sup> cells

The adenoviral vector expressing APP695 has been described previously<sup>28</sup>. Primary embryonic fibroblasts derived from mice deficient in COX-1 and COX-2 were infected with 100 plaque-forming units per cell in serum-free medium for 2 h. Medium was changed and cells were then treated with drugs as described above.

### APP processing, A $\beta$ clearance and Notch cleavage

Methods and antibodies for analysis of steady-state APP expression, APPs secretion, APP turnover and APP internalization are described in the Supplementary Information. A $\beta$  turnover after NSAID treatment was examined by culturing untransfected CHO cells with conditioned medium obtained from WT-APP-transfected CHO cells. CHO cells were pretreated with sulindac sulphide overnight. Medium was then replaced with the conditioned medium supplemented with increasing concentrations of sulindac sulphide and incubated for another 24 h. A $\beta$ 40 and A $\beta$ 42 levels in the medium were then analysed by bicine/urea A $\beta$  western blot analysis and ELISA as documented below.

The Myc-tagged, amino-terminal-truncated Notch-1 construct (NΔEMV, in which methionine 1,726 has been mutated to valine to eliminate translation initiation at that site) and the construct containing only the NICD domain have been described<sup>22</sup>. To monitor formation of NICD, the NΔEMV construct was transiently transfected into HEK293 cells. Cells were then treated with drugs for 36 h and subsequently pulse labelled with [<sup>35</sup>S]-methionine for 30 min and chased for 2 h. Cell lysates were immunoprecipitated with monoclonal antibody 9E10 (Calbiochem) against the Myc-epitope sequence, fractionated by SDS-PAGE and analysed by phosphor imaging.

## Mass spectrometry

Secreted Aβ peptides were analysed by IPMS as described<sup>29</sup>. One millilitre of conditioned medium was immunoprecipitated with monoclonal antibody 4G8 (Senetek) and molecular masses and concentrations of Aβ peptides were measured with an ultraviolet-laser desorption/ionization time-of-flight mass spectrometer. To compare the concentrations of individual Aβ species in conditioned medium, synthetic Aβ(12–28) peptide (Sigma) was added to the samples as an internal standard, and relative peak heights were calculated.

## Western blotting of Aβ peptides

Bicine/urea Aβ western blot analysis was performed as described<sup>23</sup>. One millilitre of conditioned medium was immunoprecipitated with APP monoclonal antibody 26D6 recognizing amino acids 1–12 of the Aβ sequence<sup>26</sup>. Samples were separated on bicine/urea gels, transferred to nitrocellulose membranes and immunoblotted with 26D6 antibody. Standard Aβ40, Aβ42 and Aβ38 peptides (Sigma) were separated on the same gel for identification of the corresponding Aβ species. Representative radiograms are shown.

## Ibuprofen treatment of Tg2576 transgenic mice

Female Tg2576 mice overexpressing APP695 containing the 'Swedish' mutation were treated at 3 months old, when they show high levels of soluble Aβ in brain but no signs of Aβ deposition<sup>30</sup>. NSAIDs were dissolved in DMSO, mixed with a sucrose drink (Kool-Aid) and fed orally to the animals. Controls were administered Kool-Aid with DMSO. The total amount of ibuprofen or naproxen delivered was 50 mg kg<sup>-1</sup> d<sup>-1</sup>. This daily dose was divided equally and administered every 4 h for 3 d. Two hours after the final dose, animals were killed and SDS-soluble Aβ40 and Aβ42 were analysed by ELISA as described previously<sup>30</sup>.

Received 19 July; accepted 13 September 2001.

1. Akiyama, H. *et al.* Inflammation and Alzheimer's disease. *Neurobiol. Aging* **21**, 383–421 (2000).
2. McGeer, P. L., Schulzer, M. & McGeer, E. G. Arthritis and anti-inflammatory agents as possible protective factors for Alzheimer's disease: a review of 17 epidemiologic studies. *Neurology* **47**, 425–432 (1996).
3. Rogers, J. *et al.* Clinical trial of indomethacin in Alzheimer's disease. *Neurology* **43**, 1609–1611 (1993).
4. Anthony, J. C. *et al.* Reduced prevalence of AD in users of NSAIDs and H2 receptor antagonists: the Cache County study. *Neurology* **54**, 2066–2071 (2000).
5. Stewart, W. F., Kawas, C., Corrada, M. & Metter, E. J. Risk of Alzheimer's disease and duration of NSAID use. *Neurology* **48**, 626–632 (1997).
6. Smith, W. L., DeWitt, D. L. & Garavito, R. M. Cyclooxygenases: structural, cellular, and molecular biology. *Annu. Rev. Biochem.* **69**, 145–182 (2000).
7. Younkin, S. G. The role of Aβ42 in Alzheimer's disease. *J. Physiol. Paris* **92**, 289–292 (1998).
8. Murphy, M. P. *et al.* Presenilin 1 regulates pharmacologically distinct γ-secretase activities. Implications for the role of presenilin in γ-secretase cleavage. *J. Biol. Chem.* **275**, 26277–26284 (2000).
9. McEvoy, G. K. *AHFS Drug Information* (American Society of Health-System Pharmacists, Bethesda, 1998).
10. Janne, P. A. & Mayer, R. J. Chemoprevention of colorectal cancer. *N. Engl. J. Med.* **342**, 1960–1968 (2000).
11. Lim, G. P. *et al.* Ibuprofen suppresses plaque pathology and inflammation in a mouse model for Alzheimer's disease. *J. Neurosci.* **20**, 5709–5714 (2000).
12. Piazza, G. A. *et al.* Apoptosis primarily accounts for the growth-inhibitory properties of sulindac metabolites and involves a mechanism that is independent of cyclooxygenase inhibition, cell cycle arrest, and p53 induction. *Cancer Res.* **57**, 2452–2459 (1997).
13. Cryer, B. & Feldman, M. Cyclooxygenase-1 and cyclooxygenase-2 selectivity of widely used nonsteroidal anti-inflammatory drugs. *Am. J. Med.* **104**, 413–421 (1998).
14. Zhang, X., Morham, S. G., Langenbach, R. & Young, D. A. Malignant transformation and antineoplastic actions of nonsteroidal antiinflammatory drugs (NSAIDs) on cyclooxygenase-null embryo fibroblasts. *J. Exp. Med.* **190**, 451–459 (1999).
15. Abdel-Halim, M. S., Sjoquist, B. & Anggard, E. Inhibition of prostaglandin synthesis in rat brain. *Acta Pharmacol. Toxicol.* **43**, 266–272 (1978).
16. Li, Y. M. *et al.* Photoactivated γ-secretase inhibitors directed to the active site covalently label presenilin 1. *Nature* **405**, 689–694 (2000).
17. De Strooper, B. *et al.* A presenilin-1-dependent γ-secretase-like protease mediates release of Notch intracellular domain. *Nature* **398**, 518–522 (1999).
18. Hadland, B. K. *et al.* γ-secretase inhibitors repress thymocyte development. *Proc. Natl Acad. Sci. USA* **98**, 7487–7491 (2001).
19. Yankner, B. A. *et al.* Neurotoxicity of a fragment of the amyloid precursor associated with Alzheimer's disease. *Science* **245**, 417–420 (1989).
20. Oster-Granite, M. L., McPhie, D. L., Greenan, J. & Neve, R. L. Age-dependent neuronal and synaptic degeneration in mice transgenic for the C terminus of the amyloid precursor protein. *J. Neurosci.* **16**, 6732–6741 (1996).
21. Perez, R. G. *et al.* Mutagenesis identifies new signals for β-amyloid precursor protein endocytosis,

turnover, and the generation of secreted fragments, including Aβ42. *J. Biol. Chem.* **274**, 18851–18856 (1999).

22. Kopan, R., Schroeter, E. H., Weintraub, H. & Nye, J. S. Signal transduction by activated mNotch: importance of proteolytic processing and its regulation by the extracellular domain. *Proc. Natl Acad. Sci. USA* **93**, 1683–1688 (1996).
23. Wiltfang, J. *et al.* Improved electrophoretic separation and immunoblotting of β-amyloid (Aβ) peptides 1–40, 1–42, and 1–43. *Electrophoresis* **18**, 527–532 (1997).
24. Qiu, W. Q. *et al.* Insulin-degrading enzyme regulates extracellular levels of amyloid β-protein by degradation. *J. Biol. Chem.* **273**, 32730–32738 (1998).
25. Iwata, N. *et al.* Identification of the major Aβ<sub>1–42</sub>-degrading catabolic pathway in brain parenchyma: suppression leads to biochemical and pathological deposition. *Nature Med.* **6**, 143–150 (2000).
26. Kang, D. E. *et al.* Modulation of amyloid β-protein clearance and Alzheimer's disease susceptibility by the LDL receptor-related protein pathway. *J. Clin. Invest.* **106**, 1159–1166 (2000).
27. McGeer, P. L. Cyclo-oxygenase-2 inhibitors: rationale and therapeutic potential for Alzheimer's disease. *Drugs Aging* **17**, 1–11 (2000).
28. Yuan, H. *et al.* Recombinant adenovirus is an appropriate vector for endocytotic protein trafficking studies in cultured neurons. *J. Neurosci. Methods* **88**, 45–54 (1999).
29. Wang, R., Sweeney, D., Gandy, S. E. & Sisodia, S. S. The profile of soluble amyloid β protein in cultured cell media. Detection and quantification of amyloid β protein and variants by immunoprecipitation-mass spectrometry. *J. Biol. Chem.* **271**, 31894–31902 (1996).
30. Kawarabayashi, T. *et al.* Age-dependent changes in brain, CSF, and plasma amyloid (β) protein in the Tg2576 transgenic mouse model of Alzheimer's disease. *J. Neurosci.* **21**, 372–381 (2001).

**Supplementary information** is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

## Acknowledgements

We thank X. Zhang and D. A. Young for fibroblasts deficient in COX-1 and COX-2; R. Kopan for Notch plasmids; M. Kounnas for 26D6 antibody; P. Sims, D. Galasko and C. Eckman for discussions; and Takeda industries for the BAN50, BA27 and BC05 antibodies. We also thank P. Needleman for his scientific input. This work was supported by National Institutes of Health grants (E.H.K., T.E.G., R.W.), a Beeson Award (T.E.G.), an Ellison Medical Foundation New Scholars award (T.E.G.), a Robert and Clarice Smith Fellowship (P.D.), a John Douglas French Alzheimer's Foundation Fellowship (M.P.M.), and an Emmy Noether fellowship from the Deutsche Forschungsgemeinschaft (S.W.).

Correspondence and requests for materials should be addressed to E.H.K. (e-mail: [edkoo@ucsd.edu](mailto:edkoo@ucsd.edu)).

# Gridlock signalling pathway fashions the first embryonic artery

Tao P. Zhong\*, Sarah Childs\*, James P. Leu & Mark C. Fishman

Cardiovascular Research Center, Massachusetts General Hospital;  
Department of Medicine, Harvard Medical School, 149 13th Street,  
Charlestown, Massachusetts 02129, USA

Arteries and veins are morphologically, functionally and molecularly very different, but how this distinction is established during vasculogenesis is unknown<sup>1,2</sup>. Here we show, by lineage tracking in zebrafish embryos, that angioblast precursors for the trunk artery and vein are spatially mixed in the lateral posterior mesoderm. Progeny of each angioblast, however, are restricted to one of the vessels. This arterial–venous decision is guided by *gridlock* (*grl*), an artery-restricted gene that is expressed in the lateral posterior mesoderm<sup>3</sup>. Graded reduction of *grl* expression, by mutation or morpholino antisense, progressively ablates regions of the artery, and expands contiguous regions of the vein, preceded by an increase in expression of the venous marker EphB4 receptor (*ephb4*)<sup>2</sup> and diminution of expression of the arterial marker ephrin-B2 (*efnb2*)<sup>2</sup>. *grl* is downstream of *notch*<sup>4</sup>, and interference with *notch* signalling, by blocking *Su(H)*<sup>4</sup>, similarly reduces the artery and increases the vein. Thus, a *notch–grl* pathway controls assembly of the first embryonic artery, apparently by adjudicating an arterial versus venous cell fate decision.

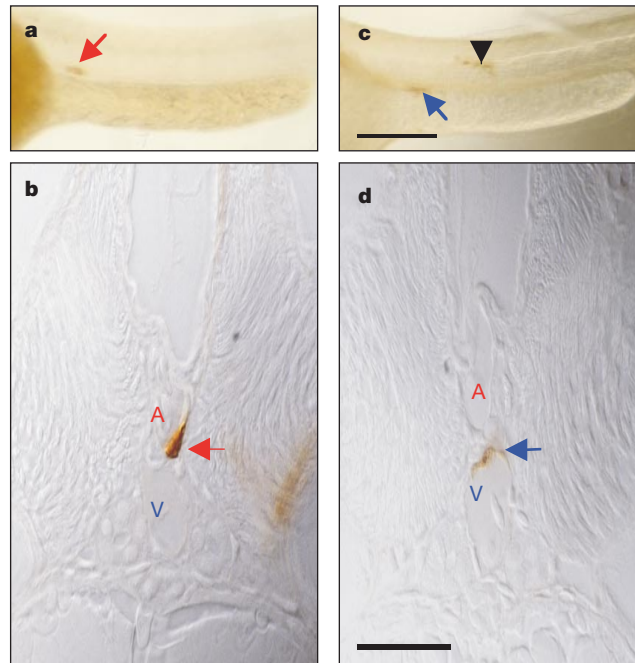
\* Present addresses: Department of Medicine and Cell Biology, Vanderbilt Medical School, Nashville, Tennessee 37232, USA (T.P.Z.); and Department of Biochemistry and Molecular Biology, University of Calgary, Calgary, Alberta, T2N 4N1, Canada (S.C.).



The first artery and vein of the vertebrate embryo are formed in the trunk, and when the two are joined they comprise the original cardiovascular loop. These simple endothelial tubes are formed by *de novo* aggregation of free angioblasts, a process termed

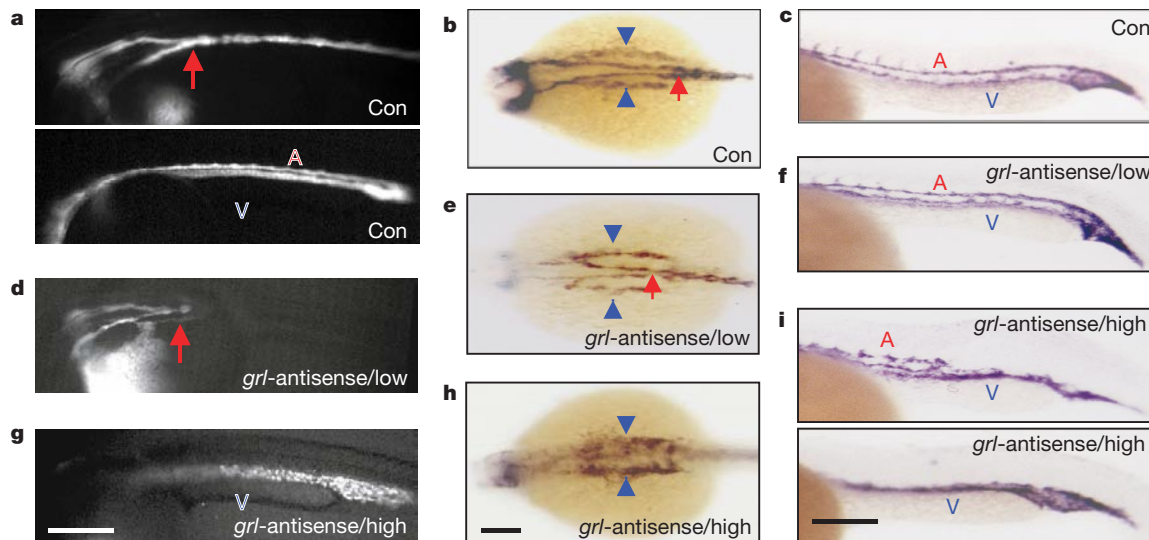
vasculogenesis (as opposed to angiogenesis, the sprouting from pre-existing vessels).

In zebrafish, the midline trunk artery and vein are different from the time of their first formation. Although both vessels express *flt5*



**Figure 1** Tracking the lineage of pre-arterial and pre-venous angioblasts from the lateral posterior mesoderm to the trunk vessels. **a, b**, In some experiments progeny are detected in the aorta (A and red arrow). Whole-mount immunostaining in lateral view (**a**) and histological cross-section (**b**) at 40 h after fertilization. **c, d**, In other experiments, progeny

are found in the vein (V and blue arrow). Whole-mount immunostaining in lateral view (periderm (black arrow) labelling along the track of the laser is readily distinguished) (**c**), and cross-section (**d**). Dorsal is up and anterior is to the left (**a, c**). Scale bars: **a, c**, 100  $\mu$ m; **b, d**, 50  $\mu$ m.



**Figure 2** Reduction in *Grl* produces graded loss of regions of the aorta and expansion of veins. **a–c**, Control embryos. **a**, Top, angiogram in dorsal view showing the aortic bifurcation (arrow); bottom, lateral view angiogram showing the aorta (A) dorsal to the vein (V) in the trunk of the embryo. **b**, Dorsal view of anterior region of the embryo by *flk in situ* labelling showing the anterior aortic bifurcation (arrow) and the anterior cardinal veins laterally (arrowheads). **c**, Trunk region—by *flk in situ* labelling—showing the aorta (A) and posterior cardinal vein (V). **d–f**, Embryos injected with a low dose of morpholino *grl*-antisense oligonucleotides (*grl*-antisense; ~3.5 ng) phenocopy the *grl* mutation, with disruption of the anterior aortic bifurcation. **d**, Dorsal view of angiogram showing the point of functional blockage to circulation at the anterior aortic bifurcation (arrow). **e**, *flk* labelling showing the point of disruption near the aortic bifurcation, corresponding to the point of

functional blockage (arrow). Arrowheads indicate the cardinal veins. **f**, The vessels of the trunk, the aorta (A) and posterior cardinal vein (V) appear normal. **g–i**, Embryos injected with a high dose of morpholino *grl*-antisense oligonucleotides (~12 ng) additionally lose large sections, or all, of the aorta. **g**, Angiogram in lateral view showing filling of the trunk vein only, with little or no arterial system. **h**, *flk* labelling shows no evident anterior aortae. The cardinal veins (arrowheads) are expanded. **i**, In the trunk region there is loss of the posterior regions of the aorta (top) or loss of the entire aorta (bottom), with increased *flk* labelling of the vein. For a low dose of morpholino *grl*-antisense oligonucleotides (**d–f**), the anterior defect is seen in 78% ( $n = 56$ ) of injections; for high a dose (**g–i**), the additional trunk defect is seen in 26% ( $n = 49$ ) of injections. Dorsal is up and anterior is to the left (**a, c, d, f, g, i**). Scale bars, 100  $\mu$ m.

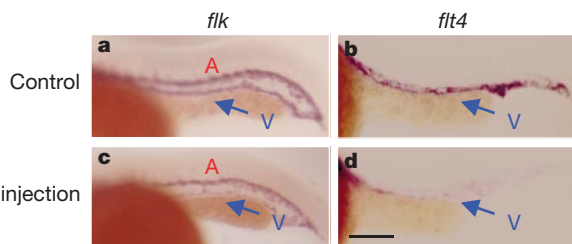
and the VEGF receptor, *flk*<sup>6</sup>, *grl* is expressed only in the artery<sup>3</sup> and *flt4* only in the vein<sup>5</sup> at 24 h after fertilization. Before vessel formation all four of these genes are expressed in the lateral posterior mesoderm (LPM), and, as development proceeds, in progressively more medial positions. This is compatible with DiI labelling of cell patches in frog<sup>7</sup> and chick/quail transplantation<sup>8</sup>, suggesting that the LPM contains precursors for these vessels.

We examined the source of the vasculogenic angioblasts by labelling cells in the lateral mesoderm with fluorescent dextran and then tracking their migration. The cells are too small for direct injection; therefore, we injected one–two-cell embryos with 4,5-dimethoxy-2-nitrobenzyl (DMNB)-caged fluorescent dextran, and later activated (uncaged) the dye in a few cells of the LPM during somitogenesis using a laser<sup>9</sup>. In some cases, progeny of these LPM cells generate short segments of tubes running in the midline along the anterior–posterior axis. These can be readily distinguished as artery or vein (Fig. 1a, c), depending on dorsal–ventral position. The LPM contains precursors for both the artery and vein<sup>10</sup>, but they seem to be rare. Of 158 uncaging experiments, only 82 appeared to have progeny in either the artery or vein. In all cases, only the artery (red arrow in Fig. 1a, b) or vein (blue arrow in Fig. 1c, d) was labelled (32 artery; 50 vein). It is clear that the vascular fate of an angioblast is determined in the LPM. Arterial and venous precursors are spatially mixed within LPM, and an individual angioblast contributes to only one or the other of these vessels.

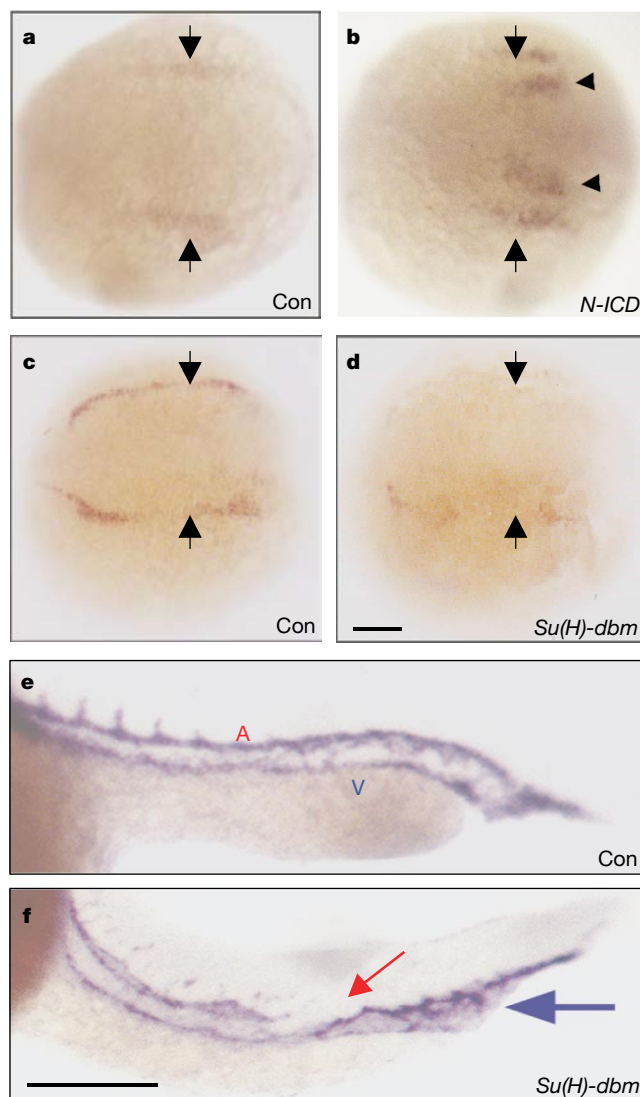
Because *grl* is expressed in the LPM and then along the length of the aorta<sup>3</sup>, we wondered whether it might have a role in the angioblast cell fate decision. The *grl* mutation is hypomorphic and interrupts blood flow by causing a blockage at the anterior bifurcation of the dorsal aorta<sup>11</sup>. To better assay the cellular basis of the defect, we examined the pattern of *flk* expression in that region and found that in *grl* mutant embryos there is a small interruption of the vessel limited to this anterior bifurcation point. This suggested that the anterior bifurcation is the site that is most sensitive to the reduced levels of *grl* caused by the mutation. To examine the effects of further diminution in *grl*, we injected embryos with *grl*-antisense oligonucleotides that were modified by morpholino, and analysed the functional integrity of the circulation by microangiography (Fig. 2a, d, g). We also examined the cellular constituency of the vessels by *flk* *in situ* hybridization (Fig. 2b, c shows the anterior bifurcation and the trunk region, respectively). Morpholino oligonucleotides are believed to block binding of small ribosomal subunits to RNA and thereby reduce protein synthesis<sup>12</sup>. To assay this effect, we developed antibodies to predicted Grl peptides, and used doses of morpholino *grl*-antisense for injections that could reduce Grl protein levels by western analysis (data not shown).

Injection of embryos with these morpholino *grl*-antisense constructs causes a dose-related disruption of increasingly large segments of the aorta. The most sensitive region is the anterior

bifurcation, so that the lowest effective doses create a blockage to blood flow at the anterior aortic bifurcation (Fig. 2d), exactly in the region affected by the *grl* mutation<sup>11</sup>. There is disruption of *flk* labelling in the anterior bifurcation (arrow in Fig. 2e) but the trunk vasculature appears normal (Fig. 2f). With injection of higher doses of *grl*-antisense oligonucleotides there is no visible arterial system, even anteriorly. Microangiographic dye, normally carried forward by flow through the aorta, instead retrogradely fills the posterior cardinal vein (Fig. 2g). *flk* *in situ* examination reveals that the arterial system is eliminated, including the aortic bifurcation (Fig. 2h) and the entire dorsal aorta (Fig. 2i). In some cases there is a preserved remnant of the rostral part of dorsal aorta (Fig. 2i).



**Figure 3** Injection of *grl* into wild-type embryos reduces the vein. **a, b**, In control embryos artery and vein both express *flk* (**a**), but only the vein expresses *flt4* (**b**). **c**, Injection of *grl* (~200 pg) essentially eliminates all *flk* labelling in the vein without noticeably affecting the artery (51%, *n* = 52). **d**, *flt4* expression is eliminated by overexpression of *grl* (~200 pg), confirming that the affected vessel is the vein (53%, *n* = 44). Lateral views and anterior is to the left. A, artery; V, vein. Scale bar, 100 μm.



**Figure 4** Expression of *grl* is increased by injection of activated *notch1* and blocked by the dominant negative inhibitor of *Su(H)*. **a**, Expression of *grl* (arrows) in lateral posterior mesoderm at the 5-somite stage in a control embryo. **b**, After injection of human *Notch1-ICD*<sup>29</sup> (~200 pg) there is expansion of *grl* expression into presomitic mesoderm (arrowheads) (83%, *n* = 34). Injection of *Xenopus notch1-ICD*<sup>30</sup> caused the same effects as human *Notch1-ICD* (data not shown). **c**, *grl* mRNA (arrows) in the lateral mesoderm at the 3-somite stage in a control embryo. **d**, Injection of *Su(H)-dbm* (~350 pg) reduces *grl* expression at this stage (35%, *n* = 46). **e**, *flk* labelling of the trunk (lateral view) at the 25-somite stage shows the artery and vein in a control embryo. **f**, After injection of *Su(H)-dbm* (~350 pg) there is loss of the posterior segment of the artery (red arrow) and expansion of the adjacent region of the vein (blue arrow) at the 25-somite stage (25%, *n* = 45). Dorsal views and anterior is to the left (**a–d**); lateral views and anterior is to the left (**e** and **f**). A, artery; V, vein. Scale bars, 100 μm.

Where the aorta is diminished, the vein appears to be expanded. This venous expansion is most prominent when the artery is completely absent (Fig. 2h, i). This demonstrates that *grl* is required for artery formation along its entire length, and indicates that vein is expanded at the expense of artery.

The expansion of vein when Grl is removed suggests that Grl may

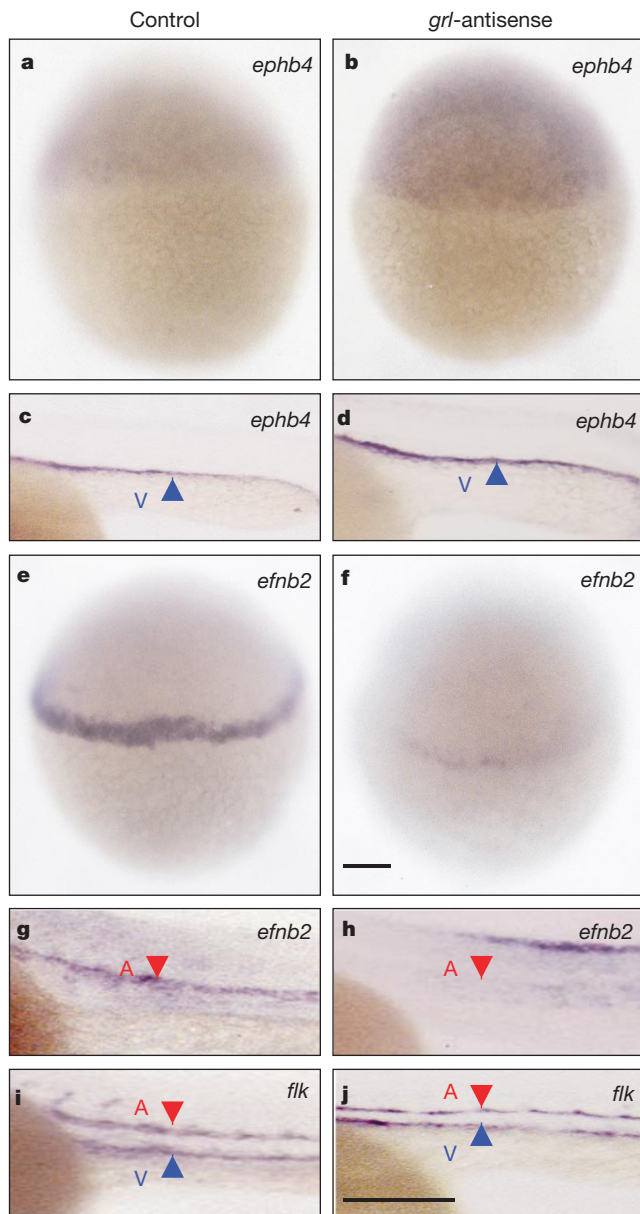
act to suppress the venous fate. To examine this issue directly, we overexpressed *grl* messenger RNA by injection into one–two-cell embryos. High levels of *grl* cause a marked diminution in the vein, revealed by *flk* (Fig. 3a, c) and *flk4* (Fig. 3b, d) *in situ* labelling. Overexpression of *grl* does not seem to affect aortic size, suggesting that *grl* is necessary but not sufficient for formation of artery.

Grl is a member of a family of proteins that are related to Hes proteins termed hairy-related transcription factors, Hrt<sup>13</sup> (also known as Hey<sup>14</sup>, Hsr<sup>15</sup> and Chf<sup>16</sup>). The *hes* genes are transcriptional repressors with several domains, including a characteristic basic helix-loop-helix motif and a carboxy-terminal WRPW motif; the latter is critical for the binding of corepressors<sup>17</sup>. Grl/Hrt proteins contain a similar motif organization, but have distinctive residues in the bHLH domain and a YRPW motif near the C terminus instead of the WRPW motif. *hes* genes are essential to many cell fate decisions, including the sorting out of cells from groups of apparently equivalent progenitor cells<sup>18</sup>. In these cases the *hes* genes often function downstream of the cell surface Notch receptor to suppress a particular cell fate. Recent work indicates that *hrt* genes in the mouse can also be regulated by the *notch* pathway at the transcriptional level<sup>19</sup>. We therefore examined whether *grl* expression can be increased by Notch activation. In zebrafish and other species there are many *notch* genes, several of which are expressed in the LPM and early vessels<sup>20,21</sup>. Introduction of the Notch1 intracellular domain (Notch1-ICD) serves to mimic activation of any of the *notch* genes<sup>4</sup>. Injection of *notch1-ICD* induces *grl* expression in the presomitic mesoderm at the 5-somite stage (Fig. 4a, b). This supports the suggestion that *grl* expression is downstream of *notch*. It does not demonstrate that *grl* normally is regulated by this pathway.

Suppressor of hairless (*Su(H)*) is an intermediary in the Notch signalling pathway, acting by binding to sites upstream of *hes* genes<sup>22</sup>. We found a conserved *Su(H)/rbp-1*-binding site (CGTGGGAA) between –47 and –55 bp upstream of the ATG of the *grl* gene. We examined whether interference with *Su(H)* diminishes *grl* expression and function. *Su(H)* activity can be blocked by a construct (*Su(H)-dbm*) designed to associate with the intracellular domain of Notch and interfere with DNA binding<sup>23</sup>. Injection of this dominant negative *Su(H)* construct reduces *grl* expression in the lateral mesoderm at the 3-somite stage (Fig. 4c, d), confirming that *grl* is a downstream transcriptional target of *Su(H)*. Increasing the dose of dominant negative *Su(H)* phenocopies graded loss of Grl. At the lowest doses it causes selective disruption of the anterior aortic bifurcation, as occurs in the *grl* mutation. As doses of dominant negative *Su(H)* are increased, there is loss of progressively longer segments of the aorta (the most posterior being the most sensitive), revealed by *flk* *in situ* labelling (Fig. 4e, f). Where the aorta is absent, the vein is expanded.

If *grl* acts to repress particular fates, similar to proteins of the Hes family<sup>4,18</sup>, downstream effectors might include markers of the venous fate. One candidate is the *ephb4* receptor<sup>24</sup>, which we found is venous-specific in zebrafish (Fig. 5c), as it is in mice<sup>25</sup>. Before vessel formation, *ephb4* is not detectable in the LPM; however, it is detectable earlier in a diffuse pattern throughout the blastoderm (Fig. 5a), as is *grl* at this stage (data not shown). Reduction in Grl by low-dose morpholino *grl*-antisense oligonucleotides increases *ephb4* expression at this stage (Fig. 5b). Once the vessels have formed, there is increased *ephb4* expression in the vein (Fig. 5c, d). As in mouse, *ephb4* (refs 25, 26) is an arterial marker in zebrafish (Fig. 5g). Its expression is diminished when Grl is reduced at both early and late stages (Fig. 5e–h) at doses of antisense for which both trunk vessels are still present (Fig. 5i, j). This suggests that normally *ephb4* is repressed as a downstream venous effector in responding to the *grl* signal, whereas expression of its reciprocal arterial gene *efnb2* is enhanced.

Thus, the *notch*–*grl* pathway regulates vasculogenesis. In particular, it is necessary for normal artery formation. The different regions of the aorta exhibit different sensitivities to loss of function



**Figure 5** Reduction of Grl causes an increase of *ephb4* expression and decrease of *ephb2* expression. **a**, *ephb4* is expressed in a diffuse pattern at the 50% epiboly stage (middle of gastrulation) in a control embryo. **b**, Injections of a low dose of morpholino *grl*-antisense oligonucleotides (~3.5 ng) increase *ephb4* expression at the 50% epiboly stage (97%,  $n = 30$ ). **c**, At the 30-somite stage, *ephb4* RNA is expressed in the vein in a control embryo. **d**, The increased *ephb4* RNA caused by the injection of a low dose of morpholino *grl*-antisense is restricted to the vein (64%,  $n = 28$ ). **e**, *efnb2* is restricted to the germ ring at the 50% epiboly stage in a control embryo. **f**, Injections of low-dose morpholino *grl*-antisense oligonucleotides (~3.5 ng) decrease *efnb2* expression (78%,  $n = 32$ ). **g**, At the 30-somite stage, *efnb2* RNA is restricted to the artery in a control embryo. **h**, The decreased *efnb2* RNA caused by the injection of morpholino *grl*-antisense is diminished in the artery (56%,  $n = 25$ ). This low dose of morpholino *grl*-antisense was selected because it does not abolish the artery as revealed by *flk* staining (**i** and **j**). Dorsal views with anterior at the top (**a**, **b**, **e**, **f**); lateral views with anterior to the left (**c**, **d**, **g**–**j**). Artery (A) is indicated by a red arrowhead; vein (V) is indicated by a blue arrowhead. Scale bars, 100  $\mu$ m.



in this pathway, with the anterior bifurcation and the posterior artery the most sensitive, and the anterior trunk the least. This may reflect positional information along the anterior–posterior axis, with use of additional or complementary components of the Notch signalling system at different levels. Interestingly, targeted mutation of *notch1* in mice selectively disrupts the anterior bifurcation of the aorta<sup>27</sup>.

One model for how this regulation might occur, compatible with evidence here and with known features of Notch activity in other cells<sup>4</sup>, is that the *notch–grr* pathway regulates a choice of cell fates, in this case as artery or vein. Overexpression of *grr* diminishes the vein without increasing the artery, so it does not seem to act instructively to generate the artery. Because blockade of the *notch–grr* pathway diminishes artery and, additionally, increases the contiguous vein, we suggest that the normal action of *grr* is to repress the venous fate, which is mediated through downstream effectors including *ephb4* and possibly others. Such repression might be necessary for the arterial fate to emerge. This is compatible with other cell fate decisions, where assumption of a primary fate by one precursor cell represses an identical fate in neighbouring cells by activating in them the *notch–hes* pathway<sup>18,28</sup>. This could explain why angioblast clones labelled in the LPM populate only artery or vein.

Clearly there is an essential and early developmental distinction between artery and vein, in addition to those differences later imposed by blood pressure. These attributes may be relevant to understanding of vessel specificity of diseases and to the design of synthetic conduits for vascular repair. □

## Methods

### Laser-activated lineage tracking

Zebrafish embryos were injected at the one–two-cell stage with 1% DMNB-caged fluorescein (Molecular Probes) in 160 mM KCl, and grown overnight in the dark at 24 °C. Between the 7- and 12-somite stages of development, a patch of 5–10 cells in the lateral plate mesoderm was activated by multiple pulses of a nitrogen laser tuned to 365 nm (Laser Sciences), using a Micropoint Laser Ablation System (Photronics) mounted on a Zeiss Axioplan microscope. The degree of activation was assessed immediately by epifluorescence. The embryos were grown in the dark at 28 °C overnight and fixed in 4% PFA/PBT at the equivalent of 40 h of normal development at 28 °C. Embryos were permeabilized in cold methanol/acetone. To block the endogenous peroxidase staining of blood, embryos were incubated with 0.5% H<sub>2</sub>O<sub>2</sub> in 0.5% sheep serum for 30 min at room temperature before antibody staining. Uncaged fluorescein was detected with a 1/500 dilution of an anti-fluorescein antibody that was conjugated with peroxidase (Boehringer Mannheim), and visualized with diaminobenzidine.

### Microangiography, whole-mount *in situ* labelling and histology

Microangiography was performed as described<sup>11</sup> with the exception that 0.2-μm yellow-green polystyrene Fluospheres (Molecular Probes) in 1% BSA were sonicated for 2 min at full power before injection into the sinus venosus. We carried out whole-mount RNA *in situ* hybridization for *grr*, *flk*, *flr* and *flt4* as described<sup>3</sup>. For histological analysis, specimens were fixed in 4% paraformaldehyde, dehydrated and embedded in plastic (JB-4). Nomarski photomicroscopy was performed with an Axiophot using Ektachrome 160T film (Zeiss). Wild M5 and M10 dissecting microscopes equipped with Nikon cameras were used for low-power photomicroscopy.

### RNA and morpholino oligonucleotide injection

Sense-capped RNA was synthesized using the mMESSAGE mMACHINE system (Ambion) from the following linearized plasmids: *pCS2+/Su(H)-dbm* (*Xenopus* dominant negative mutant form of *Su(H)*, digested with *NotI*, transcribed with Sp6); *pCDNA3/cytoTANI* (human intracellular domain of Notch1; *XhoI* digestion and T7 transcription); *pCRII-TOPO/grr* (zebrafish *Grr*, digested by *HindIII*, transcribed by T7). RNA was diluted at ~40 μg ml<sup>-1</sup> in solution A (0.1% phenol red, 0.2 M KCl) and microinjected into the blastomeres at the one–four-cell stage of embryos that had been dechorionated by pronase treatment. Siblings from the same pool were injected using solution A as the controls. The volume of injected RNA was determined by measuring the diameter of a spherical drop of RNA injected into mineral oil on an objective micrometer with 0.01-mm divisions.

Morpholino-modified antisense oligonucleotides (5'-CGCGCAGGTACA GACACCAA AACT-3') to the *grr* gene were designed to target the 5' untranslated region containing the binding site of 40S ribosomal subunit with no predicted internal hairpin structures. The control oligonucleotide, designed by sequence scrambling of *grr*-antisense oligonucleotides, and an unrelated morpholino oligonucleotide (provided by Gene Tools, LLC) were injected as controls. Neither of these controls causes a vessel defect. The morpholino oligonucleotides were solubilized in 1× Danieau solution (58 mM NaCl, 0.7 mM KCl, 0.4 mM MgSO<sub>4</sub>, 5 mM HEPES, pH 7.6) at the concentration of 1 mM (~8 mg ml<sup>-1</sup>).

For injection of sense RNA and morpholino antisense oligonucleotide, different doses were tested for determining the effective dose required for causing the specific vascular phenotype and the maximal dose for causing the detrimental effects. We carried out injections using Microinjector 5242.

Received 14 August; accepted 26 September 2001.

- Carmeliet, P. Mechanisms of angiogenesis and arteriogenesis. *Nature Med.* **6**, 389–395 (2000).
- Yancopoulos, G. D., Klagsbrun, M. & Folkman, J. Vasculogenesis, angiogenesis, and growth factors: ephrins enter the fray at the border. *Cell* **93**, 661–664 (1998).
- Zhong, T. P., Rosenberg, M., Mohideen, M. A., Weinstein, B. & Fishman, M. C. Gridlock, an HLH gene required for assembly of the aorta in zebrafish. *Science* **287**, 1820–1824 (2000).
- Artavanis-Tsakonas, S., Rand, M. D. & Lake, R. J. Notch signaling: cell fate control and signal integration in development. *Science* **284**, 770–776 (1999).
- Thompson, M. A. *et al.* The cloche and spadetail genes differentially affect hematopoiesis and vasculogenesis. *Dev. Biol.* **197**, 248–269 (1998).
- Fouquet, B., Weinstein, B. M., Serluca, F. C. & Fishman, M. C. Vessel patterning in the embryo of the zebrafish: guidance by notochord. *Dev. Biol.* **183**, 37–48 (1997).
- Cleaver, O. & Krieg, P. A. VEGF mediates angioblast migration during development of the dorsal aorta in *Xenopus*. *Development* **125**, 3905–3914 (1998).
- Pardanaud, L. & Dieterlen-Lievre, F. Manipulation of the angiopoietic/hemangiopoietic commitment in the avian embryo. *Development* **126**, 617–627 (1999).
- Serbedzija, G. N., Chen, J. N. & Fishman, M. C. Regulation in the heart field of zebrafish. *Development* **125**, 1095–1101 (1998).
- Noden, D. M. Embryonic origins and assembly of blood vessels. *Am. Rev. Respir. Dis.* **140**, 1097–1103 (1989).
- Weinstein, B. M., Stemple, D. L., Driever, W. & Fishman, M. C. Gridlock, a localized heritable vascular patterning defect in the zebrafish. *Nature Med.* **1**, 1143–1147 (1995).
- Nasevicius, A. & Ekker, S. C. Effective targeted gene 'knockdown' in zebrafish. *Nature Genet.* **26**, 216–220 (2000).
- Nakagawa, O., Nakagawa, M., Richardson, J. A., Olson, E. N. & Srivastava, D. HRT1, HRT2, and HRT3: a new subclass of bHLH transcription factors marking specific cardiac, somitic, and pharyngeal arch segments. *Dev. Biol.* **216**, 72–84 (1999).
- Leimeister, C., Externbrink, A., Klamt, B. & Gessler, M. *Hey* genes: a novel subfamily of hairy- and Enhancer of split-related genes specifically expressed during mouse embryogenesis. *Mech. Dev.* **85**, 173–177 (1999).
- Kokubo, H., Lun, Y. & Johnson, R. L. Identification and expression of a novel family of bHLH cDNAs related to *Drosophila* hairy and enhancer of split. *Biochem. Biophys. Res. Commun.* **260**, 459–465 (1999).
- Chin, M. T. *et al.* Cardiovascular basic helix loop helix factor 1, a novel transcriptional repressor expressed preferentially in the developing and adult cardiovascular system. *J. Biol. Chem.* **275**, 6381–6387 (2000).
- Fisher, A. L., Ohsaka, S. & Caudy, M. The WRPW motif of the hairy-related basic helix-loop-helix repressor proteins acts as a 4-amino-acid transcription repression and protein–protein interaction domain. *Mol. Cell Biol.* **16**, 2670–2677 (1996).
- Fisher, A. & Caudy, M. The function of hairy-related bHLH repressor proteins in cell fate decisions. *Bioessays* **20**, 298–306 (1998).
- Nakagawa, O. *et al.* Members of the HRT family of basic helix-loop-helix proteins act as transcriptional repressors downstream of Notch signaling. *Proc. Natl Acad. Sci. USA* **97**, 13655–13660 (2000).
- Bierkamp, C. & Campos-Ortega, J. A. A zebrafish homologue of the *Drosophila* neurogenic gene Notch and its pattern of transcription during early embryogenesis. *Mech. Dev.* **43**, 87–100 (1993).
- Shirayoshi, Y. *et al.* Proto-oncogene of int-3, a mouse Notch homologue, is expressed in endothelial cells during early embryogenesis. *Genes Cells* **2**, 213–224 (1997).
- Bailey, A. M. & Posakony, J. W. Suppressor of hairless directly activates transcription of enhancer of split complex genes in response to Notch receptor activity. *Genes Dev.* **9**, 2609–2622 (1995).
- Wettestein, D. A., Turner, D. L. & Kintner, C. The *Xenopus* homolog of *Drosophila* Suppressor of Hairless mediates Notch signaling during primary neurogenesis. *Development* **124**, 693–702 (1997).
- Holder, N. & Klein, R. Eph receptors and ephrins: effectors of morphogenesis. *Development* **126**, 2033–2044 (1999).
- Wang, H. U., Chen, Z. F. & Anderson, D. J. Molecular distinction and angiogenic interaction between embryonic arteries and veins revealed by ephrin-B2 and its receptor Eph-B4. *Cell* **93**, 741–753 (1998).
- Chan, J. *et al.* Morphogenesis of prechordal plate and notochord requires intact Eph/ephrin B signaling. *Dev. Biol.* **234**, 470–482 (2001).
- Krebs, L. T. *et al.* Notch signaling is essential for vascular morphogenesis in mice. *Genes Dev.* **14**, 1343–1352 (2000).
- Jensen, J. *et al.* Control of endodermal endocrine development by Hes-1. *Nature Genet.* **24**, 36–44 (2000).
- Sestan, N., Artavanis-Tsakonas, S. & Rakic, P. Contact-dependent inhibition of cortical neurite growth mediated by notch signaling. *Science* **286**, 741–746 (1999).
- Furukawa, T. *et al.* Rax, Hes1, and Notch1 promote the formation of Muller glia by postnatal retinal progenitor cells. *Neuron* **26**, 383–394 (2000).

## Acknowledgements

We thank members of the Fishman laboratory for advice during these experiments, especially J.-N. Chen for help with the preliminary work on lineage analysis and C. Simpson for histological analyses. We thank S. Artavanis-Tsakonas and M. Rand for human Notch1-ICD construct; C. Kintner for *Xenopus* notch-ICD and Su(H)-Dbm constructs; and J. Chen for zebrafish ephrin-B2 and EphB4 probes. J.P.L. was supported by the Sarnoff Foundation. This work is supported by grants from the National Institutes of Health and a sponsored research agreement with Genentech to M.C.F.

Correspondence and requests for materials should be addressed to M.C.F. (e-mail: fishman@civrc.harvard.edu).

# Asynchronous replication and allelic exclusion in the immune system

Raul Mostoslavsky<sup>\*†</sup>, Nandita Singh<sup>\*‡</sup>, Toyoaki Tenzen<sup>†§</sup>, Maya Goldmit<sup>†</sup>, Chana Gabay<sup>†</sup>, Sharon Elizur<sup>†</sup>, Peimin Qi<sup>‡</sup>, Benjamin E. Reubinoff<sup>||</sup>, Andrew Chess<sup>‡¶</sup>, Howard Cedar<sup>†</sup> & Yehudit Bergman<sup>†</sup>

<sup>†</sup>Departments of Cellular Biochemistry & Human Genetics, and Experimental Medicine & Cancer Research, PO Box 12272, Hebrew University, Jerusalem 91120, Israel

<sup>‡</sup>Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA

<sup>§</sup>Department of Evolutionary Genetics, National Institute of Genetics, Mishima, Shizuoka-Ken, 411-8540, Japan

<sup>||</sup>Department of Obstetrics and Gynecology & The Goldyn Savad Institute of Gene Therapy, Hadassah Ein-Kerem University Hospital, Jerusalem 91120, Israel

<sup>¶</sup>Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

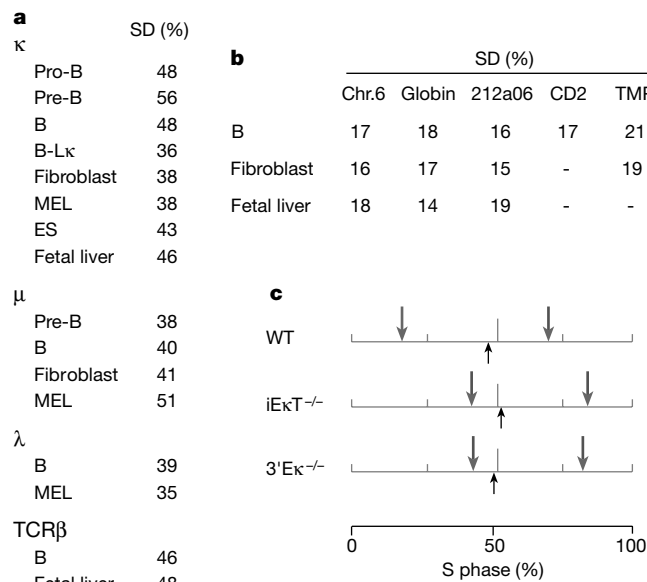
\* These authors contributed equally to this work

The development of mature B cells involves a series of molecular decisions which culminate in the expression of a single light-chain and heavy-chain antigen receptor on the cell surface<sup>1,2</sup>. There are two alleles for each receptor locus, so the ultimate choice of one receptor type must involve a process of allelic exclusion. One way to do this is with a feedback mechanism that downregulates rearrangement after the generation of a productive receptor molecule<sup>3</sup>, but recent work suggests that monoallelic epigenetic changes may also take place even before rearrangement<sup>4</sup>. To better

understand the basis for distinguishing between alleles, we have analysed DNA replication timing. Here we show that all of the B-cell-receptor loci ( $\mu$ ,  $\kappa$  and  $\lambda$ ) and the TCR $\beta$  locus replicate asynchronously. This pattern, which is established randomly in each cell early in development and maintained by cloning, represents an epigenetic mark for allelic exclusion, because it is almost always the early-replicating allele which is initially selected to undergo rearrangement in B cells. These results indicate that allelic exclusion in the immune system may be very similar to the process of X chromosome inactivation.

The entire genome is divided into a series of well defined zones which replicate in a programmed manner throughout the S phase of the cell cycle, and this pattern is correlated with gene expression. Housekeeping genes all replicate in early S phase, whereas many tissue-specific genes are developmentally regulated to replicate late in most tissues, but early in the cells that express these genes<sup>5</sup>. Replication timing is easily assayed by fluorescence *in situ* hybridization (FISH) analysis of interphase nuclei, which reveals a single dot for an as-yet-unreplicated locus and a double signal after replication<sup>6</sup>. In a non-synchronized population of cells, a high percentage of double signals is indicative of early replication, and a preponderance of single dots is characteristic of late-replicating DNA. Monoallelically expressed sequences, such as imprinted genes<sup>7</sup>, olfactory receptors<sup>8</sup> or the X chromosome in females<sup>9</sup>, all show an asynchronous replication pattern whereby one allele replicates earlier than the other in each cell. This is easily detected by the presence of a high percentage of cells showing one single and one double hybridization signal in the FISH assay<sup>7</sup>.

Because antigen receptor genes are expressed monoallelically in the lymphoid lineage, we examined these loci for replication timing patterns using FISH. Initial studies were carried out for the  $\kappa$  light-chain gene in normal B cells. Strikingly, over 40% of interphase



**Figure 1** Antigen receptor loci replicate asynchronously. **a**, Cell lines or stimulated B and T cells were labelled with BrdU and then prepared and assayed for replication timing using probes for the antigen receptor loci. For each experiment the number of cells with single/single (SS), double/double (DD) or single/double (SD) signals were recorded, but only the %SD is shown in the chart. In each case 100–300 nuclei were counted. In another experiment (not shown), probes for C $\kappa$  alone (35% SD), C $\kappa$  together with the most 3' V $\kappa$  sequences (36% SD), V $\kappa$  located 2 Mb upstream (36% SD) and R $\beta$  located 40 kb downstream (35% SD) all showed similar asynchronous replication, indicating that this pattern is regional, and the same was true for C $\mu$  (35% SD) and a probe located about 170 kb downstream of C $\mu$  (37% SD) at the heavy-chain locus. B-L $\kappa$  are B cells from a mouse carrying a transgenic  $\kappa$  locus<sup>15</sup>. The unrearranged endogenous alleles were

assayed using the R $\beta$  probe, which represents a region not included in the transgene. **b**, The %SD for several control probes is shown. Additional controls show synchronous replication (<20%SD) for erythroleukaemia (MEL) and F9 (which is similar to embryonic stem cells, ES); see ref. 7. **c**, Replication timing was measured at the  $\kappa$  locus (using C $\kappa$ ) in B cells from a wild-type (WT) mouse and from mice carrying homozygous targeted deletions for the intronic (iE $\kappa$ T<sup>-/-</sup>) or the 3' (3'E $\kappa$ T<sup>-/-</sup>)  $\kappa$  enhancers. The results are shown graphically, with the first arrow (from the left) representing the percentage of SS cells and the second arrow representing the percentage of SS + SD cells. This presentation allows us to visualize (upper, thick arrows) the time in S phase when the first and second alleles replicate. The average percentage of single signals for a control probe (CD2) is shown (lower, thin arrows) for each cell type.

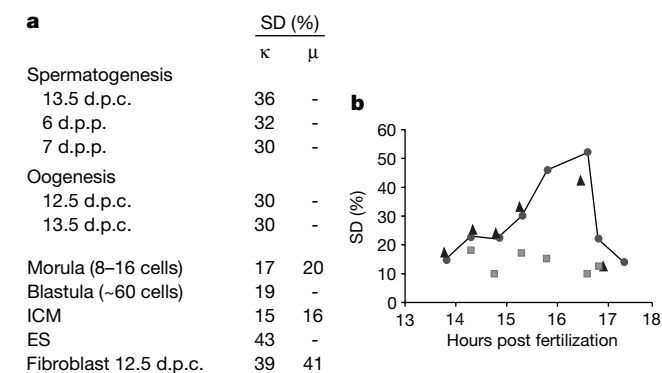
nuclei show a single/double pattern (Fig. 1a), indicating that the two alleles at this locus indeed replicate asynchronously, in contrast to a number of different control genes which show a low level of single/double signals (Fig. 1b). A similar pattern was observed in proliferating cells from earlier stages of B-cell development and even in non-lymphoid cells types, such as erythroleukaemia or primary fibroblasts, suggesting that the pattern must be set up early in development (Fig. 1a). Indeed, unlike other structural features of the  $\kappa$  locus which are specific for B cells<sup>10</sup>, the control of this phenomenon is not affected by the major  $\kappa$  enhancers, and when they are removed by targeted deletion<sup>11,12</sup>, overall replication timing shifts to a later time, but the two alleles still undergo DNA synthesis asynchronously (Fig. 1c). We note that in addition to the  $C_{\kappa}$  probe used in these experiments, probes for the  $V_{\kappa}$  region (BAC 122K2) located 2 Mb upstream and the ribose 5-phosphate isomerase (*Rpi*) gene located 40 kb downstream<sup>13</sup> also showed asynchronous replication (35–36% single/double) (details in legend to Fig. 1), and similar results were obtained for the  $\mu$  locus (see legend to Fig. 1), indicating that this is a regional phenomenon with a pattern very similar to that observed for imprinted gene domains<sup>7</sup>. Furthermore, this asynchronous replication profile is not unique to the  $\kappa$  locus, and similar data were obtained for the B-cell-specific  $\lambda$  light chain and  $\mu$  heavy chain<sup>31</sup>, as well as the T-cell-associated TCR- $\beta$  locus (Fig. 1a).

Asynchronous replication of antigen receptor loci is not restricted to the lymphoid system and can be observed in a variety of different cell types, so we next asked when this pattern is set up during development. Analysis of replication timing in sequential stages of spermatogenesis and oogenesis demonstrated that the two  $\kappa$  alleles replicate asynchronously before meiosis (Fig. 2a). This differential marking is evidently retained in mature germ cells as well, since the same pattern was observed again immediately after fertilization during the first division cycle in zygotes (Fig. 2b). This pattern is clearly lost, however, in the morula and blastula, where both alleles

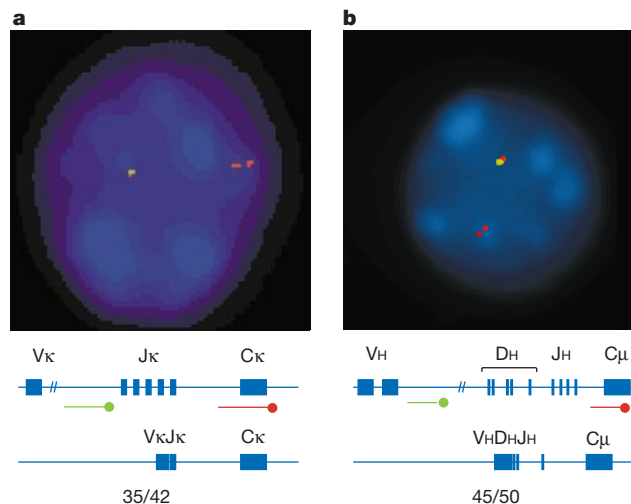
replicate relatively synchronously (about 20% single/double), and this erasure was even observed in purified inner cell mass (ICM) cells, which actually serve as the source for the embryo proper. Asynchronous replication was again observed in embryonic stem cells (which mimic the post-blastula embryo), and in early fibroblasts. A similar phenomenon was observed for the  $\mu$  heavy-chain locus, as well (Fig. 2a, b). This pattern is distinctly different from those of imprinted genes: such genes are programmed in the gametes to replicate in a way that is specific to the parent of origin; they then maintain this asynchronous pattern through pre-implantation development<sup>14</sup> (see legend to Fig. 2a). In contrast, differential replication timing at the antigen receptor loci is erased in the morula and then re-established, probably at about the time of implantation. The monoallelically expressed olfactory receptor genes and IL-4 behave in a similar manner<sup>14</sup>.

The establishment of the observed asynchrony of replication well before B-cell development suggests that it could provide a differential mark, rendering one allele more accessible for rearrangement. With this in mind, we wanted to test whether asynchronous replication timing is itself correlated with the selection of a single allele for rearrangement. Thus we devised an *in situ* system capable of distinguishing between the germline and rearranged copies of the  $\kappa$  locus. This was done by using two separate FISH probes, one for  $C_{\kappa}$ , and a second which hybridizes upstream of the  $J_{\kappa}$  segments in a region normally deleted as part of the rearrangement process. An analysis of over 40 B-cell nuclei exhibiting an asynchronous replication pattern (single/double) for the  $C_{\kappa}$  probe shows that in 83% of the cases, the late-replicating copy (single dot) is associated with the germline (non-deleted) allele (Fig. 3a). These findings indicate that rearrangement in B cells almost always occurs on the early-replicating allele, and this appears to be a general phenomenon for the antigen receptor loci, since analysis of the  $\mu$  locus gave almost identical results (90% of cases) (Fig. 3b).

The asynchronous replication timing pattern that we observe appears to be independent of antigen receptor rearrangement and



**Figure 2** Asynchronous replication during development. **a**, Replication timing analysis of the  $\kappa$  locus in cells for various stages of gametogenesis and in the early embryo, including purified ICM. As a control, the imprinted gene *Snrpn* was also tested and found to be asynchronously replicating in ICM cells (33% SD). In each case 100–200 nuclei were scored for SS, SD and DD signals but only the percentage SD is shown in the chart. Results using control probes for cells in gametogenesis and for pre-implantation embryos (morula and blastula) have been published<sup>14</sup> and shown to be statistically different from probes demonstrating asynchronous replication. **b**, A mixture of zygotes from 15–20 superovulated females was incubated and assayed for SD signals by FISH at the indicated time on the day after mating using  $\kappa$  (circles),  $\mu$  (triangles) and control (squares) probes. Since all of the zygotes undergo their first round of replication at about the same time, we see a peak for the SD pattern (16–17 h post fertilization). Premeiotic germ cells presumably have one early and one late allele in each cell. Thus, zygotes could be formed with either two early, two late or one early and one late allele, implying that only 50% of the cells should show asynchrony at the peak time. We note that for imprinted genes, where all of the zygotes carry asynchronous alleles, the maximum level of SD signals was much higher<sup>14</sup>. d.p.c., days post coitum; d.p.p. days post partum.



**Figure 3** Rearrangement is associated with early replication. **a**, B cells were assayed by FISH using a probe for  $C_{\kappa}$  (pSP1g8) (red) and a second probe for the region upstream to  $J_{\kappa}1$  (pGL1.6) (green) which is deleted in most rearranged cells. Nuclei showing a single/double signal for  $C_{\kappa}$  were scored for the position of the intact germline allele in the same cell. In 35 of 42 cells, the germ line allele was associated with the single signal. It should be noted that in the mouse, rearrangement can also occur by inversion, in which case the upstream  $J_{\kappa}1$  probe is not deleted, but rather shifted to a distal position. Cells of this nature were not included in the survey. **b**, A similar experiment was done using a  $C_{\mu}$  probe (red) and a second probe upstream to the D region (green). In 45 of 50 cells, the allele which did not undergo V to DJ rearrangement is associated with the single, unreplicated signal.



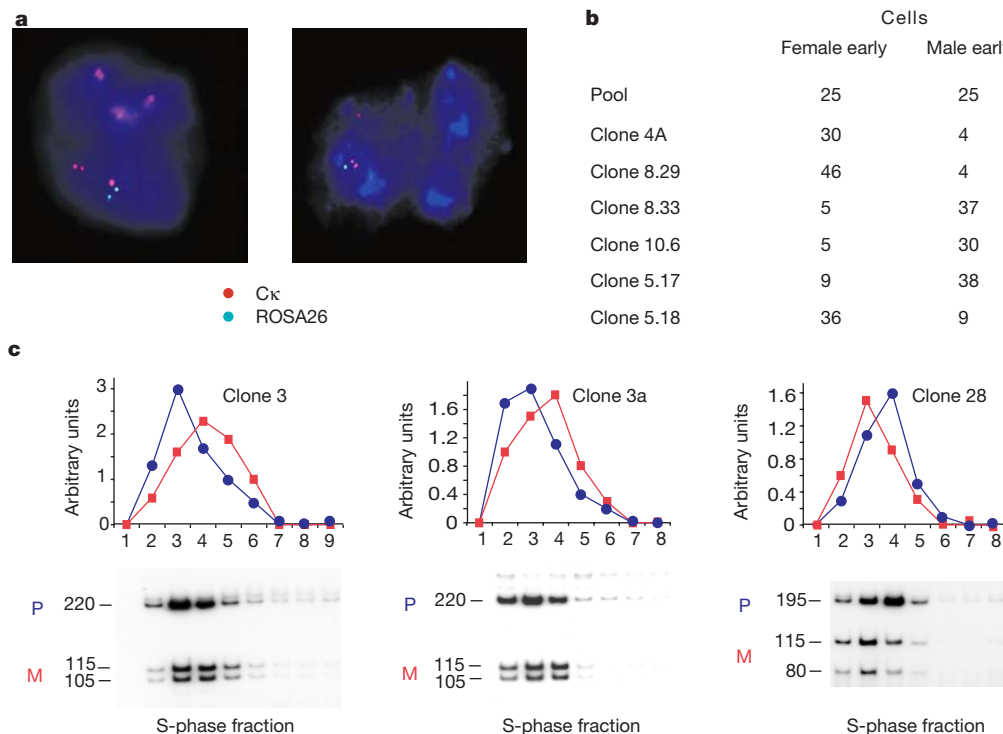
expression. In spleen cells carrying a  $\kappa$  transgene (B-L $\kappa$ ), for example, both endogenous  $\kappa$  loci are in the silent germline configuration and not expressed<sup>15</sup>, yet these alleles still replicate asynchronously (36% single/double) (see Fig. 1a). We also analysed replication timing in B cells from knock-in mice carrying a pre-rearranged  $\kappa$  gene on both alleles. These B cells express both alleles at a high level (R.M., Y.B. H.C. & K. Rajewsky, unpublished results), yet still have asynchronous replication of the  $\kappa$  alleles (36% single/double).

Because the initial rearrangement at the  $\kappa$  locus can occur randomly on either the maternal or paternal allele in each cell<sup>16</sup>, we assumed that early-replication timing must also be set up in a stochastic manner: some cells would have an early-replicating maternal allele and others an early-replicating paternal allele. To test this hypothesis, we generated primary fibroblasts from a mouse that was heterozygous for an insertion of *neo-lacZ* into the ROSA26 locus on chromosome 6 (ref. 17), not far (approximately 18 centimorgans) from the  $\kappa$  locus; we then used this marker to identify the parental alleles in each cell. By double-label FISH we scored cells containing a single/double pattern for C $\kappa$ . In 50% of these nuclei, the paternal allele replicated early, while in the remainder of the nuclei the maternal allele did so. We then asked whether these patterns, once established, are maintained in a clonal manner. Fibroblasts were transformed by infection with a retrovirus encoding SV-40 large T antigen and individual clones were then expanded. We analysed a number of these clones for  $\kappa$  replication timing using the *neo-lacZ* probe to identify the parental origin of the alleles. Each clone demonstrated a consistent pattern of allele-specific early replication. In clones 4A, 8.29 and 5.18, for example,

the ROSA26 transgene was almost always associated with the double signal in cells showing an asynchronous pattern, whereas in clones 8.33, 10.6 and 5.17, the pattern was reversed (Fig. 4a, b).

To confirm these results, we used the alternate approach of cell-cycle fractionation to measure replication timing. We generated individual pre-B-cell clones transformed by infection with Abelson virus from *Mus spretus*  $\times$  *M. musculus* F<sub>1</sub> mice. We labelled replicating DNA with bromodeoxyuridine (BrdU) and then used FACS (fluorescence-activated cell sorting) to separate the cells into seven to eight S-phase fractions. BrdU DNA was then isolated from each fraction and assayed for sequence content by quantitative polymerase chain reaction (PCR)<sup>18</sup>, taking advantage of a restriction site polymorphism in the  $\kappa$  region to identify the two parental alleles. As shown in Fig. 4c, each individual clone displays a clear asynchronous replication pattern, with one allele replicating before the other. Out of five clones analysed by this technique, three had the paternal *M. spretus* allele replicating early (for example, clone 3), and two clones had the maternal *M. musculus* allele first to replicate (for example, clone 28). These patterns remained stable for over 50 generations in culture and even after subcloning (for example, clone 3a).

These studies clarify the mechanism of allelic exclusion in the immune system. Unlike models based on the equal availability of both alleles to rearrangement<sup>3</sup>, our results suggest that asynchronous replication timing represents a fundamental epigenetic difference between the two alleles in each cell and that this serves as the basis for initially choosing one copy to undergo demethylation and subsequent rearrangement. It is probably this feature, in fact, which keeps the remaining germline allele in a closed con-



**Figure 4**  $\kappa$  asynchronous replication is heritable. **a**, Embryonic fibroblasts from mice carrying the *neo-lacZ* marked maternal ROSA26 transgene were assayed by FISH using the C $\kappa$  (red) and a *neo-lacZ* probe (green) for the ROSA26 transgene. C $\kappa$  single/doubles were then scored for the location of ROSA26. In 50% (25/50) of the cells, the single was associated with the ROSA26-carrying maternal chromosome 6 (left) and in 50% the single was associated with the paternal allele (right). **b**, Pooled fibroblasts and individual ear fibroblast subclones originating from four mice (numbered 4, 5, 8 and 10) were assayed as above for the number of paternal early or maternal early nuclei among cells showing an asynchronous replication pattern. **c**, Replication timing was analysed by S-phase

fractionation of BrdU-labelled Abelson-virus-transformed pre-B cells (see Methods) isolated from *M. spretus*  $\times$  *M. musculus* F<sub>1</sub> mice. The  $\kappa$  locus was amplified by PCR and *HhaI* then used to distinguish between the paternal (blue) *M. spretus* allele (uncut) and the maternal (red) *M. musculus* allele (cut). For clones 3 and 3a we used primers 170-1 and K2 which yield a 220-bp product, while clone 28 was assayed using primers 170-1 and K4 (195-bp product). We repeated each of these experiments at least three times. PCR analysis of a control region (mouse globin) from clone 3 showed a completely synchronous replication pattern (data not shown).

formation and thus provides enough time for a productive antigen receptor molecule to interact with the cell surface and ultimately bring about feedback inhibition of the rearrangement machinery. In the absence of a productive receptor, we assume that given enough time, the second, later-replicating allele can also undergo both demethylation and rearrangement<sup>4</sup>. According to this model, receptor loci continue to retain their differential epigenetic marks even after the cells have already decided on a single expressing allele. This may help explain why further alterations as a result of antigen-induced receptor editing<sup>19</sup> or somatic hypermutation<sup>20</sup> appear to be restricted to the already rearranged alleles. We have not investigated how the replication timing pattern affects rearrangement, but it is likely that this mechanism is indirect, probably involving chromatin structure. In this sense, it may well combine with many other components which control the cell specificity and timing of rearrangement through DNA accessibility<sup>10,32</sup>.

The differential marking of two alleles in the same cell, as occurs for the antigen receptor loci, is reminiscent of the process of X-chromosome inactivation. In both cases the regions involved replicate asynchronously in all somatic cells<sup>9</sup>. At the local level, the  $\kappa$  gene itself behaves in a manner analogous to X-chromosome-linked tissue-specific genes, which also appear to undergo monoallelic demethylation and transcriptional activation, but only in the cell type of expression<sup>21,22</sup>. Developmentally, these processes are also quite similar. Just as in the X chromosomes, differences between the alleles at the antigen receptor loci are erased in the morula<sup>23</sup>, and asynchronous replication is then re-established at about the time of implantation<sup>9</sup>. In both cases, the choice of which allele replicates early is random, with either the maternal or paternal copy being 'picked'; once established, replication asynchrony is clonally inherited through future cell divisions. Furthermore, it is always the early-replicating allele which is active. Thus, we have shown that this process of monoallelic inactivation is not unique to the X chromosome, but can also take place, albeit in a regional manner, on autosomes as well. Other loci, such as the olfactory receptor gene clusters<sup>8</sup> and the cytokines *Il-2* and *Il-4* (ref. 24) also replicate asynchronously in a variety of different cell types and are expressed monoallelically, suggesting that this may represent a common early embryonic mechanism for distinguishing between alleles.

**Note added in proof:** Another observation consistent with asynchronous replication of the  $\mu$  heavy chain in mature B cells (Fig. 1) has been recently reported<sup>31</sup>. □

## Methods

### Embryos

Superovulated C57Black  $\times$  BALB/C F<sub>1</sub> female mice were mated with males and zygotes were removed from the oviducts about 14–20 h after hormone injection. Zygotes were incubated at 37 °C in M16 medium (Sigma) to obtain pre-implantation embryos at different stages of development, and ICM was purified by immunosurgery<sup>25</sup>.

### Cells

Fetal livers were dissected out of 12.5 d.p.c. (days post coitum) embryos and mechanically disrupted in a solution of RPMI + 10% FCS by passing through a syringe. The remainder of the embryo was disrupted with a 2-ml syringe into a solution of DMEM + 10% FCS in order to obtain embryonic fibroblasts. These cells were grown for 1–2 h. Purity of the erythroid cells from embryonic liver (85–90%) was determined by FACS analysis using a specific monoclonal antibody. Germ cells were prepared from the gonads of 12.5 d.p.c. or 13.5 d.p.c. embryos by antibody-directed flow sorting, as described<sup>26</sup>. Primitive type A and type A + type B spermatogonia were isolated from testes of 6- and 8-day-old mice by gradient sedimentation<sup>27</sup>. The purity of germ cells in the recovered populations was consistently more than 85%.

B cells were prepared by disruption of spleens isolated from the wild type, iEκT<sup>−/−</sup> (ref. 11) and 3κE<sup>−/−</sup> (ref. 12) mice in phosphate-buffered saline (PBS) followed by gentle pipetting and centrifugation through PBS. B cells were stimulated by growing at 37 °C in RPMI medium containing 10% FCS, supplemented with 20  $\mu$ g ml<sup>−1</sup> lipopolysaccharide (LPS) for 3 days. The resulting population typically contained >90% positive B cells as demonstrated by flow cytometry analysis. Stromal cell/IL7-dependent cell lines, representing pro-B cells, were derived from fetal liver of wild-type mice, as described<sup>28</sup>. Pre-B cells were obtained from spleens of Rag<sup>−/−</sup> mice carrying a  $\mu$  transgene in which B-cell differentiation was arrested at the pre-B-cell stage<sup>29</sup>. T cells were isolated from the thymus and grown in the presence of ConA (2  $\mu$ g ml<sup>−1</sup>) for 3 days in culture.

### Fibroblast isolation and SV-40 transformation

Primary mouse ear fibroblasts were isolated from ROSA26 reporter mice<sup>37</sup>. The skin was gently peeled off individual ears to expose fibroblasts, which were rubbed vigorously on a 10-mm tissue culture dish with the aid of a sterile pipette. The fibroblasts were allowed to grow for 8–10 days. T-antigen transformation was carried out by infection with supernatant from producer cell line  $\psi$ 2-SV40 (ref. 30). Cells were passaged whenever they reached confluence, 3–4 times. Following passage, the SV40-transformed cells were subcloned by limiting dilution.

### Fluorescence *in situ* hybridization

All cells used for FISH analysis, including those taken from the gonads or from embryos, were first incubated in culture with  $3 \times 10^{-5}$  M BrdU for 1 h before collection, in order to enable us to provide a marker for identifying cells in S phase. Preimplantation embryos, as well as purified ICM, were collected and fixed to a poly-L-lysine slide (Sigma) by putting them into a drop of 0.01 N HCl, 0.1% Tween 20. In the case of zygotes, cells were placed on the slide at a density which allowed the observation of pronuclei pairs without interference from other nearby zygotes. After drying, the slide was washed, dehydrated, permeabilized and fixed. All other cells were treated with hypotonic KCl solution, fixed in methanol:acetic acid (3:1) and dropped on slides, as previously described<sup>6</sup>.

FISH was performed as described previously<sup>14</sup>. Briefly, denaturation was carried out by incubation in 70% deionized formamide,  $2 \times$  SSC at 68 °C for 2 min, and then slides were dehydrated by a series of ice-cold ethanol washes (70, 90 and 100% for 5 min each). Cosmid or plasmid DNA was labelled by nick-translation, substituting dTTP (deoxythymidine triphosphate) with bio-16-dUTP or with digoxigenin-11-dUTP (Boehringer Mannheim). For detecting the  $\kappa$  chain gene in ear fibroblasts, DNA was labelled by nick translation (Amersham) with Cy3 dCTP. The critical size range of probe molecules (smaller than 500 base pairs (bp) and preferably 150–250 bp) was achieved by empirically varying the amount of DNaseI in the nick-translation reaction. Unincorporated nucleotides were separated from the probe DNA by centrifugation through 1-ml Sephadex G-50 columns (Boehringer Mannheim). Probe DNA (10–50 ng) was mixed with cot-1 (Life Technologies) (2–3  $\mu$ g) and sufficient salmon sperm DNA to obtain a total of 10  $\mu$ g in a 10- $\mu$ l hybridization cocktail. After denaturation of the probe mixture (80–90 °C for 5 min), pre-annealing of repetitive DNA sequences was carried out for 10 min at 37 °C before application to denatured nucleic-acid specimens. Following incubation overnight and subsequent post-hybridization washes, the specimens were treated with blocking solution (3% BSA,  $4 \times$  SSC) for 10 min at 37 °C in 1% BSA,  $4 \times$  SSC and 0.1% Tween 20 and slides were then washed at room temperature three times for 3 min each in  $4 \times$  SSC and 0.1% Tween 20. Biotin-labelled probes were detected with rhodamine-conjugated avidin DCS (1:500 dilution) (Vector Laboratories) and digoxigenin-labelled probes were detected with an anti-digoxigenin antibody conjugated to fluorescein isothiocyanate (FITC; Boehringer Mannheim) (1:100 dilution). BrdU was detected by anti-BrdU antibody (NeoMarkers) (1:100), followed by either rhodamine (1:50) or aminomethylcoumarin acetic acid (AMCA; 1:20) conjugated anti-mouse antibody (Jackson Immunoresearch Laboratories). Counterstaining, where needed, was done using diaminodiphenylidole (DAPI) (200 ng ml<sup>−1</sup>) in Vector antifade solution. Amplification of the digoxigenin-labelled probes was carried out with anti-sheep antibody conjugated to FITC (Vector) and of the biotin-labelled probes with biotinylated anti-avidin (Vector). For the *neo-lacZ* sequence, amplification of the digoxigenin-labelled probe was done using three antibody layers, sheep anti-digoxigenin (Roche), FITC-labelled rabbit anti-sheep (Calbiochem) and FITC-labelled goat anti-rabbit (Roche).

Replication timing profiles (per cent single/single, single/double and double/double) were determined by counting a minimum of 35–300 BrdU-positive nuclei for each cell type. The experiment involving zygote analysis was repeated over 15 times, obtaining 15–20 zygotes per female, thus enabling us to accumulate an average of 45 nuclei for each point on the graph.

The mouse probes were donated by and purchased from: B. Van Ness (pSPiB, carrying the 12.0 kb *Bam*HI fragment encompassing the J- $\kappa$  region and its derivative pGL1.6, representing the region upstream to J $\kappa$ 1); F. Salitzky (pGLRIC $\mu$ , containing the 10.7-kb *Eco*RI fragment harbouring the C $\mu$  gene); H. Schroeder (pDFL3p8, containing a 5' 3.8-kb fragment upstream of DFL16.1 in the heavy-chain locus); J. Roess (pIg303A, containing a 6-kb fragment harbouring the CA1 region); J. Chen (pBL40.1, containing 9-kb *Eco*RI fragment located 3' to V $\beta$ 14 in the TCR $\beta$  locus); L. Jackson-Grusby (pSA $\beta$ geoftrpA, containing a 5.6-kb *Xho*I fragment carrying the  $\beta$ galneo genes); D. Ward (chr 6, a random cosmid clone from chromosome 6); P. Fraser (mouse  $\beta$ -globin gene region plasmid p $\beta$ 12g and  $\beta$ major); N. Benvenisti (15-kb *TMP* (tumour-associated membrane protein) gene plasmid); I. Simon (CD2 gene-containing plasmid); H. Zachau (F1 cosmid DNA containing the 3' V $\kappa$  region); M. Reth (11.7-kb mouse *Rpi* gene pDR9 plasmid); F. Alt (4.11  $\phi$  DNA, containing 15 kb of genomic DNA located ~170 kb 3' to the C $\mu$  gene); Resgen (BAC 122K2, harbouring sequences located 2,000 kb upstream of C $\kappa$ ); Clontech (the random BAC clone, 212a06).

### Replication timing by S-phase fractionation

Abelson-virus-transformed pre-B-cell subclones from *M. spretus*  $\times$  *M. musculus* F<sub>1</sub> mice were labelled in 75  $\mu$ M BrdU for 45 min before collection, and nuclei were then sorted for cell-cycle fractions according to DNA content<sup>18</sup>. BrdU DNA was isolated from each fraction and assayed for specific sequence content by quantitative PCR carried out in two stages<sup>14</sup>, first for 18 cycles and then again for 10 cycles (95 °C 40 s, 55 °C 40 s, 72 °C 40 s) in the presence of [ $\alpha$ -<sup>32</sup>P] dCTP using the primer 5'-CAGACTGACCTCATGTGAGA-3' (170-1) together with 5'-AATGAGCAAAAGTCTACTTACG-3' (K-2) or 5'-CAGAACCAAAAGTCACAAGTA-3' (K-4) for mouse  $\kappa$ . PCR products from the S-phase fractions were electrophoresed after cutting with *Hha*I (present only on the *M. musculus* allele) and detected by autoradiography. In order to graph each allele separately, total uncut PCR

product was first quantified on an initial gel. The proportion of maternal or paternal alleles in each lane was then determined from a second gel after cutting the PCR product. These values were obtained by scanning the autoradiographs.

Received 13 June; accepted 17 September 2001.

- Pernis, B., Chiappino, G., Kelus, A. S. & Gell, P. G. Cellular localization of immunoglobulins with different allotypic specificities in rabbit lymphoid tissues. *J. Exp. Med.* **122**, 853–876 (1965).
- Cebra, J. J., Colberg, J. E. & Dray, S. Rabbit lymphoid cells differentiated with respect to alpha-, gamma-, and mu-heavy polypeptide chains and to allotypic markers Aa1 and Aa2. *J. Exp. Med.* **123**, 547–558 (1966).
- Rajewsky, K. Clonal selection and learning in the antibody system. *Nature* **381**, 751–758 (1996).
- Mostoslavsky, R. *et al.*  $\kappa$  chain monoallelic demethylation and the establishment of allelic exclusion. *Genes Dev.* **12**, 1801–1811 (1998).
- Simon, I. & Cedar, H. in *DNA Replication in Eukaryotic Cells* (ed. DePamphilis, M. L.) 387–408 (Cold Spring Harbor Laboratory Press, New York, 1996).
- Selig, S., Okumura, K., Ward, D. C. & Cedar, H. Delineation of DNA replication time zones by fluorescence in situ hybridization. *EMBO J.* **11**, 1217–1225 (1992).
- Kitsberg, D. *et al.* Allele-specific replication timing of imprinted gene regions. *Nature* **364**, 459–463 (1993).
- Chess, A., Simon, I., Cedar, H. & Axel, R. Allelic inactivation regulates olfactory receptor gene expression. *Cell* **78**, 823–834 (1994).
- Takagi, N. Differentiation of X chromosomes in early female mouse embryos. *Exp. Cell Res.* **86**, 127–135 (1974).
- Cedar, H. & Bergman, Y. Developmental regulation of immune system gene rearrangement. *Curr. Opin. Immunol.* **11**, 64–80 (1999).
- Takeda, S. *et al.* Deletion of the immunoglobulin  $\kappa$  chain intron enhancer abolishes  $\kappa$  chain gene rearrangement in *cis* but not  $\lambda$  chain gene rearrangement in *trans*. *EMBO J.* **12**, 2329–2336 (1993).
- Gorman, J. R. *et al.* The Ig $\kappa$  3' enhancer influences the ratio of Ig $\kappa$  versus Ig $\lambda$  B lymphocytes. *Immunity* **5**, 241–252 (1996).
- Apel, T. W. *et al.* The ribose 5-phosphate isomerase-encoding gene is located immediately downstream from that encoding murine immunoglobulin kappa. *Gene* **156**, 191–197 (1995).
- Simon, I. *et al.* Asynchronous replication of imprinted genes is established in the gametes and maintained during development. *Nature* **401**, 929–932 (1999).
- Betz, A. G. *et al.* Elements regulating somatic hypermutation of an immunoglobulin kappa gene: critical role for the intron enhancer/matrix attachment region. *Cell* **77**, 239–248 (1994).
- Casellas, R. *et al.* Contribution of receptor editing to the antibody repertoire. *Science* **291**, 1541–1544 (2001).
- Soriano, P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nature Genet.* **21**, 70–71 (1999).
- Tenzen, T. *et al.* Precise switching of DNA replication timing in the GC content transition area in the human major histocompatibility complex. *Mol. Cell. Biol.* **17**, 4043–4050 (1997).
- Luning-Prak, E. & Weiger, M. Light chain replacement: a new model for antibody gene rearrangement. *J. Exp. Med.* **182**, 541–548 (1995).
- Jolly, C. & Neuberger, M. Somatic hypermutation of immunoglobulin kappa transgenes: Association of mutability with demethylation. *Immunol. Cell. Biol.* **79**, 18–22 (2001).
- Mullins, L. J., Veres, G., Caskey, C. T. & Chapman, V. Differential methylation of the ornithine carbamoyl transferase gene on active and inactive mouse X chromosomes. *Mol. Cell. Biol.* **7**, 3916–3922 (1987).
- Nadon, N., Korn, N. & DeMars, R. A-11: cell type-specific and single-active-X transcription controls of newly found gene in cultured human cells. *Somat. Cell Mol. Genet.* **14**, 541–552 (1988).
- Eggan, K. *et al.* X-Chromosome inactivation in cloned mouse embryos. *Science* **290**, 1578–1581 (2000).
- Chess, A. Expansion of the allelic exclusion principle? *Science* **279**, 2067–2068 (1998).
- Hogan, B., Bedington, R., Constantini, F. & Lacey, E. *Manipulating the Mouse Embryo. A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1994).
- McCarrey, J. R., Hsu, K. C., Eddy, E. M., Klevecz, R. R. & Bolen, J. L. Isolation of viable mouse primordial germ cells by antibody-directed flow sorting. *J. Exp. Zool.* **242**, 107–111 (1987).
- Bellve, A. R. *et al.* Spermatogenic cells of the prepubertal mouse, isolation and morphological characterization. *J. Cell Biol.* **74**, 68–85 (1977).
- Rolink, A., Kudo, A., Karasuyama, H., Kikuchi, Y. & Melchers, F. Long-term proliferating early pre-B cell lines and clones with the potential to develop to surface Ig-positive, mitogen reactive B cells in vitro and in vivo. *EMBO J.* **10**, 327–336 (1991).
- Spanopoulou, E. *et al.* Functional immunoglobulin transgenes guide ordered B-cell differentiation in Rag-1-deficient mice. *Genes Dev.* **8**, 1030–1042 (1994).
- Jat, P. S., Cepko, C. L., Mulligan, R. C. & Sharp, P. A. Recombinant retroviruses encoding simian virus 40 large T antigen and polyomavirus large and middle T antigens. *Mol. Cell. Biol.* **6**, 1204–1217 (1986).
- Skok, J. A. *et al.* Nonequivalent nuclear location of immunoglobulin alleles in B lymphocytes. *Nature Immunol.* **2**, 848–854 (2001).
- Gorman, J. R. & Alt, F. Regulation of immunoglobulin light chain isotype expression. *Adv. Immunol.* **69**, 113–181 (1998).

## Acknowledgements

We thank K. Rajewsky, F. Alt and P. Soriano for kindly providing the iE $\kappa$ T<sup>+/+</sup>, 3'E $\kappa$ T<sup>+/+</sup> and ROSA26 reporter mice; B. Van Ness, F. Sablitzky, H. Schroeder, J. Roes, J. Chen, L. Jackson-Grusby, D. Ward, P. Fraser, N. Benvenisti, D. Littman, F. Alt, M. Reth, H. Zachau and I. Simon for FISH probes; B. Tuberi for help with early embryo manipulations; Z. Werb for the antibody used to purify ICM; G. Paridis, F. Ebrahimi and J. Gribnau for advice regarding cell lines and FISH; I. Simon for discussions. This work was supported by grants from the Israel Academy of sciences (Y.B. and H.C.), German Israel Foundation (Y.B.), the NIH (A.C., Y.B. and H.C.), European Community 5th Framework Quality of Life Program (Y.B.) and the Israel Cancer Research Fund (H.C.). A.C. is a Rita Allen Foundation Scholar.

Correspondence and requests for materials should be addressed to H.C. (e-mail: cedar@md2.huji.ac.il) and A.C. (e-mail: chess@wi.mit.edu).

# Identification of the cellular receptor for anthrax toxin

Kenneth A. Bradley<sup>\*†‡</sup>, Jeremy Mogridge<sup>\*§</sup>, Michael Mourez<sup>§</sup>, R. John Collier<sup>§</sup> & John A. T. Young<sup>\*</sup>

<sup>\*</sup> McArdle Laboratory for Cancer Research, University of Wisconsin—Madison, 1400 University Avenue, Madison, Wisconsin 53706, USA

<sup>†</sup> Biological and Biomedical Sciences Graduate Program, and <sup>§</sup> Department of Microbiology and Molecular Genetics, Harvard Medical School, 200 Longwood Avenue, Boston, Massachusetts 02115, USA

<sup>‡</sup> These authors contributed equally to the work

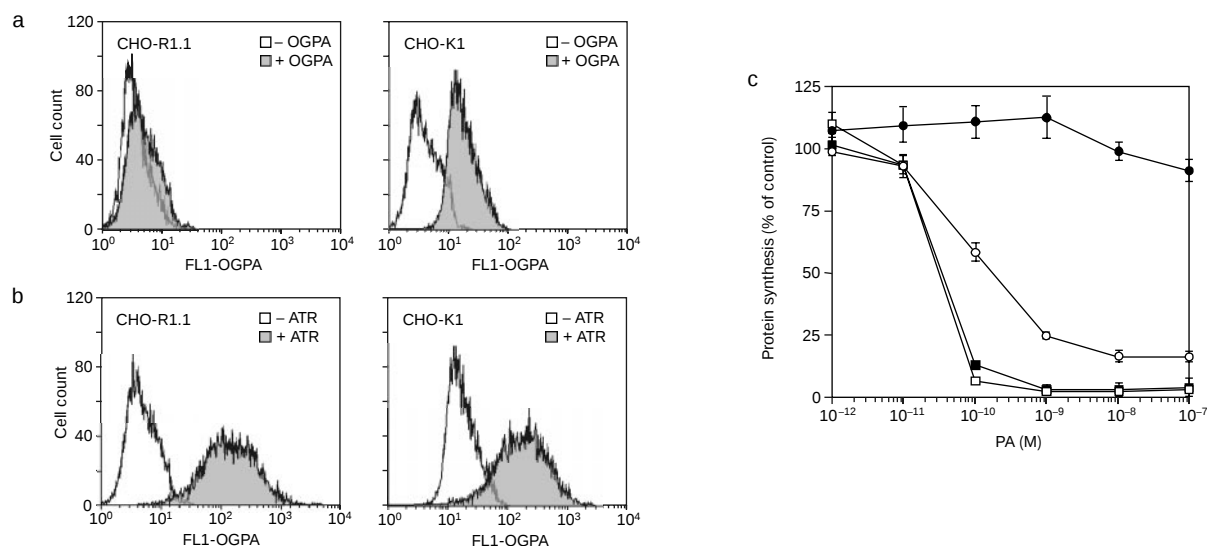
The tripartite toxin secreted by *Bacillus anthracis*, the causative agent of anthrax, helps the bacterium evade the immune system and can kill the host during a systemic infection. Two components of the toxin enzymatically modify substrates within the cytosol of mammalian cells: oedema factor (OF) is an adenylate cyclase that impairs host defences through a variety of mechanisms including inhibiting phagocytosis<sup>1,2</sup>; lethal factor (LF) is a zinc-dependent protease that cleaves mitogen-activated protein kinase and causes lysis of macrophages<sup>3–5</sup>. Protective antigen (PA), the third component, binds to a cellular receptor and mediates delivery of the enzymatic components to the cytosol. Here we describe the cloning of the human PA receptor using a genetic complementation approach. The receptor, termed ATR (anthrax toxin receptor), is a type I membrane protein with an extracellular von Willebrand factor A domain that binds directly to PA. In addition, a soluble version of this domain can protect cells from the action of the toxin.

After binding to the cell-surface receptor, PA is cleaved into two fragments by a furin-like protease<sup>6</sup>. The amino-terminal fragment, PA<sub>20</sub>, dissociates into the medium, and this allows the carboxy-terminal fragment, PA<sub>63</sub>, to heptamerize and to bind LF and OF<sup>7,8</sup>. The resulting complexes of [PA<sub>63</sub>]<sub>7</sub> with OF and/or LF are taken up into cells by receptor-mediated endocytosis and moved to a low-pH endosomal compartment<sup>9</sup>. There, the acidic environment induces a conformational change in [PA<sub>63</sub>]<sub>7</sub> that allows it to insert into the membrane and form a pore<sup>10–12</sup>. This conversion promotes the translocation of bound OF and LF across the endosomal membrane to the cytosol.

Previous studies have indicated that the receptor to which PA binds is a ubiquitous protein expressed at moderately high levels on cell surfaces (for example, 10<sup>4</sup> and 3 × 10<sup>4</sup> receptors per cell on CHO-K1 cells and macrophage cell lines, respectively)<sup>13,14</sup>. To identify this receptor, we first generated a mutant cell line lacking receptor, so that the defect could be genetically complemented. ICR-191, a DNA alkylating agent that induces small deletions and frameshift mutations in genes<sup>15</sup>, was used to introduce random mutations in the hypodiploid CHO-K1 cell line under conditions that led to about 90% cell death. The surviving mutagenized cells were then challenged with PA and LF<sub>N</sub>-DTA, a fusion protein composed of the N-terminal 255 amino acids of LF linked to the catalytic A chain of diphtheria toxin<sup>16</sup>. This recombinant toxin can kill CHO-K1 cells (in contrast to LF and PA) and it exploits the same LF-PA-receptor interactions that are required for the binding and entry of the native LF and OF proteins. Ten single-cell colonies (designated as CHO-R1.1 to CHO-R1.10) that survived toxin treatment were isolated. In control experiments performed with non-mutagenized CHO-K1 cells, no toxin-resistant cell clones were detected. One of the mutagenized clones (CHO-R1.1) was chosen for further analysis.

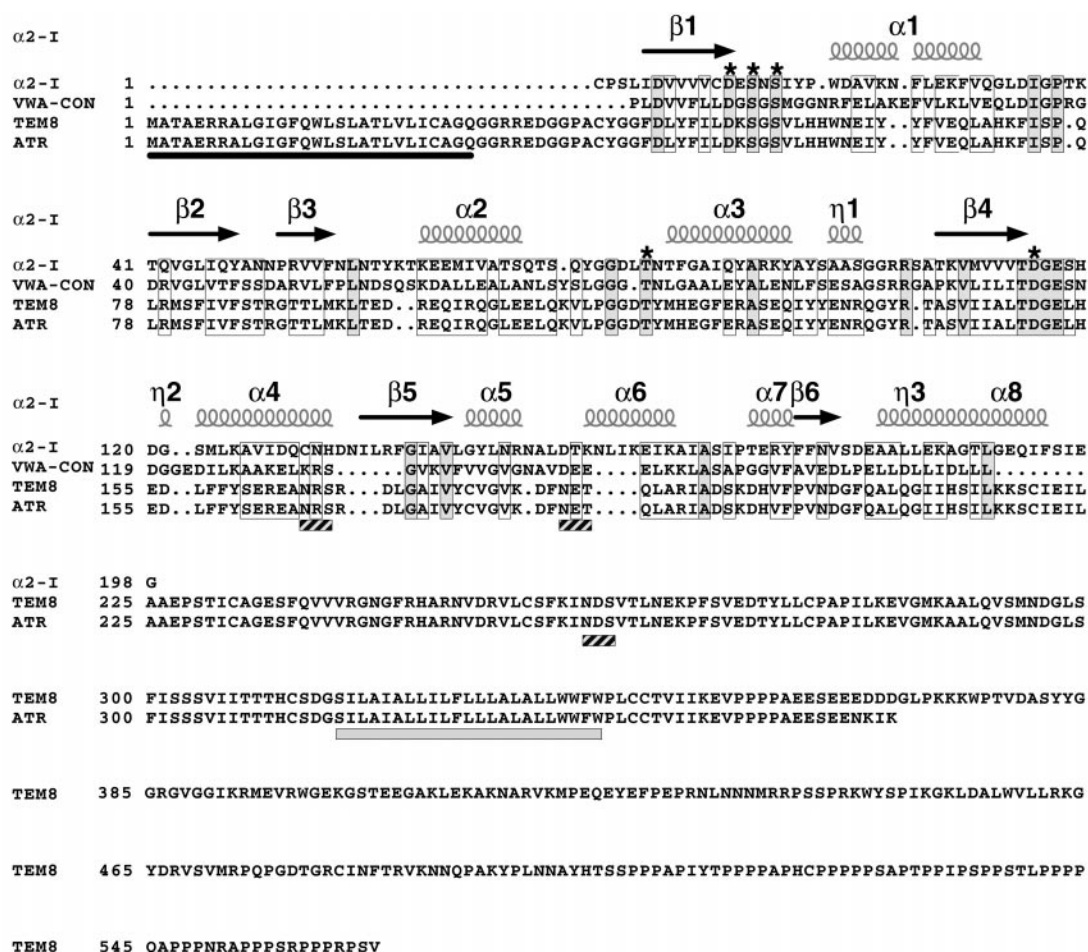
|| Present address: Department of Laboratory Medicine and Pathobiology, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada.





**Figure 1** Mutant CHO-R1.1 cells display a decreased OGPA-binding phenotype that can be corrected by overexpression of the ATR cDNA. **a**, Mutant CHO-R1.1 and wild-type CHO-K1 cells were incubated with 40 nM OGPA for 2 h on ice, washed twice then analysed by flow cytometry. **b**, Mutant and wild-type CHO cells were transduced with an MLV vector encoding ATR and then stained with OGPA as above. **c**, Expression of ATR restores toxin sensitivity. CHO-R1.1 cells (filled circles), CHO-K1 cells (filled squares) and

CHO-R1.1 and CHO-K1 cells transduced with the MLV vector encoding ATR (open circles and squares, respectively) were treated with 10<sup>-9</sup> M LF<sub>n</sub>-DTA and various concentrations of PA. About 70% of the transduced CHO-R1 cells expressed ATR as judged by OGPA staining. Medium containing 1 µCi ml<sup>-1</sup> <sup>3</sup>H-leucine was then added to cells for 1 h, and the amount of <sup>3</sup>H-leucine incorporated into cellular proteins was determined by precipitation with trichloroacetic acid and liquid scintillation counting<sup>16</sup>.



**Figure 2** Sequence alignment of ATR with the I domain of integrin α2 (α2-I), the von Willebrand factor A domain consensus sequence (VWA-CON, generated from 210 sequences aligned by the National Center for Biotechnology Information), and TEM8. The secondary structural elements are based on the crystal structure of the α2-I domain<sup>30</sup>. Conserved amino acids are boxed and identical amino acids are indicated by shaded

boxes. The putative signal sequence is underlined. The five residues that form the MIDAS motif are indicated with asterisks. The putative transmembrane domains of ATR and TEM8 are indicated by a shaded box. Potential N-linked glycosylation sites in ATR and TEM8 are indicated by hatched boxes. The alignment was made with the programs ClustalW and ESPript 1.9 (the Risler matrix was used with a global score of 0.7).

CHO-R1.1 cells were fully susceptible to killing by diphtheria toxin (data not shown), thus ruling out the possibility that resistance to PA with  $LF_N$ -DTA was due to a defect in the pathway of DT action. To test directly whether CHO-R1.1 cells lacked the receptor, we performed flow cytometric analysis using an Oregon Green-conjugated form of PA (OGPA). CHO-R1.1 cells were significantly impaired in their ability to bind to OGPA compared with the parental cell line (Fig. 1a), suggesting that these mutagenized cells had lost expression of the putative PA receptor gene. Similar analysis of the other nine mutant CHO-R1 clones demonstrated that they were also defective in binding to OGPA (data not shown).

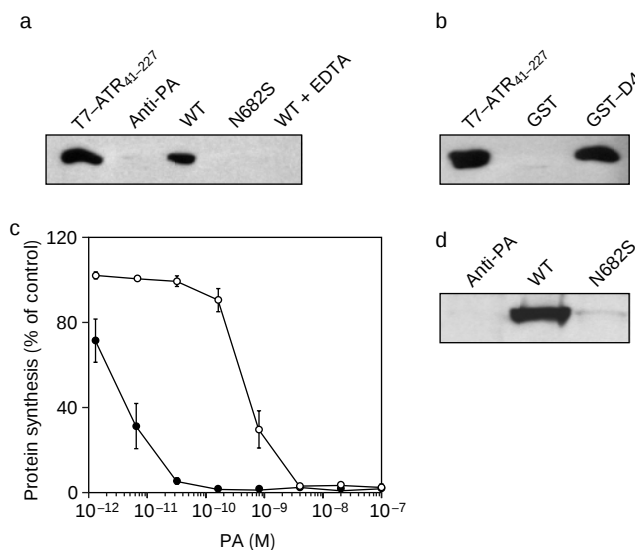
In an attempt to complement the PA-binding defect of CHO-R1.1 cells, the cells were transduced with a retrovirus-based complementary DNA library (Clontech) prepared from human HeLa cells that express the PA receptor (J.M., unpublished data). This cDNA library is contained in a murine leukemia virus (MLV) vector that is packaged into pseudotyped virus particles (MLV[VSV-G]) containing the broad host-range G protein of vesicular stomatitis virus (VSV-G)<sup>17</sup>. Retrovirus-based cDNA libraries are useful for genetic complementation approaches because they can be used to deliver a limited number of stably expressed cDNA molecules per cell. These molecules can be rapidly re-isolated by polymerase chain reaction (PCR) amplification using MLV vector-specific oligonucleotide primers<sup>18,19</sup>.

The transduced CHO-R1.1 cells were subjected to five rounds of flow cytometric sorting to isolate those that contained the cDNA clone of the putative PA receptor. Cells were sorted on the basis of their binding of OGPA in combination with an anti-PA polyclonal serum and an allophycocyanin (APC) conjugated secondary antibody. This led to the isolation of a cell population in which greater than 90% of the cells bound OGPA. This complemented cell population contained at least seven unique cDNA inserts that were obtained by the PCR amplification method described above. Each cDNA was gel purified, subcloned back into the parent pLIB vector and packaged into MLV(VSV-G) virions so that it could be tested for its ability to complement the PA-binding defect of CHO-R1.1 cells. One cDNA clone of about 1.5 kilobases (kb) (designated

as *ATR*) restored PA binding to CHO-R1.1 cells (Fig. 1b). This clone also markedly enhanced the binding of PA to parental CHO-K1 cells (Fig. 1b). Furthermore, the *ATR* cDNA clone fully restored the sensitivity of CHO-R1.1 cells to the toxin  $LF_N$ -DTA with PA (Fig. 1c).

Sequencing of the *ATR* cDNA clone revealed a single long open reading frame, encoding a 368-amino-acid protein. The protein is predicted to have a signal peptide 27 amino acids long, an extracellular domain 293 amino acids long with three putative N-linked glycosylation sites, a putative transmembrane region 23 amino acids long, and a short cytoplasmic tail (Fig. 2). A BLAST search revealed that the first 364 amino acids of *ATR* are identical to a protein encoded by the human *TEM8* cDNA clone (GenBank accession number NM032208). *TEM8* is upregulated in colorectal cancer endothelium, but the function of this protein has not been reported<sup>20</sup>. The C-terminal ends of *ATR* and the *TEM8* protein then diverge, presumably as a consequence of alternative splicing, such that *ATR* has a cytoplasmic tail of only 25 amino acids whereas *TEM8* is predicted to have a cytoplasmic tail 221 amino acids long (Fig. 2).

The most notable feature of *ATR* is the presence of an extracellular von Willebrand factor type A (VWA) domain, located between residues 44 and 216 (Fig. 2). VWA domains are present in the extracellular regions of a variety of cell surface proteins, including matrilins and integrins (designated as I domains). These domains are important for protein-protein interactions and constitute ligand-binding sites for integrins<sup>21</sup>. Ligand binding through I domains requires an intact MIDAS (metal ion-dependent adhesion site) motif<sup>22</sup>, which seems to be conserved in *ATR* (Fig. 2). The cytoplasmic tail of *ATR* contains an acidic cluster (AC motif; EESEE) that is similar to a motif found in the cytoplasmic tail of furin that specifies basolateral sorting of this protease in polarized epithelial cells<sup>23</sup>. This may be significant because the PA receptor localizes to the basolateral surface of polarized epithelial cells<sup>24</sup> and we expect that the receptor and the protease needed to bind and activate PA would be colocalized to allow for efficient entry of anthrax toxins.



**Figure 3** The VWA/I domain of *ATR* binds directly to PA. **a**, Wild-type PA (WT) or a receptor-binding mutant of PA (N682S) were mixed with T7-ATR<sub>41-227</sub> on ice for 30 min in the presence or absence of 2 mM EDTA, as indicated. A polyclonal serum specific for PA and protein A sepharose were then added, the PA-associated proteins were precipitated, subjected to SDS-PAGE, transferred to nitrocellulose, and probed with anti-T7 antibody conjugated to horseradish peroxidase. **b**, GST or GST-D4 (GST fused to domain 4 of PA) coupled to glutathione sepharose (Pharmacia) was incubated with T7-ATR<sub>41-227</sub> for 1 h at 4 °C and the samples were precipitated and analysed as described

above. **c**, CHO-K1 cells were incubated with  $LF_N$ -DTA ( $10^{-9}$  M) and various concentrations of wild-type PA (filled circles) or PA-N682S mutant (open circles) and cell viability was determined as in Fig. 1c. **d**, CHO-K1 cells were incubated with  $2 \times 10^{-8}$  M trypsin-nicked PA (wild type or N682S) for 1 h. Cells were washed with PBS, resuspended in SDS sample buffer and run on a 4–20% SDS-PAGE gel. PA was visualized by western blotting. Lane 1 (**a,b**), T7-ATR<sub>41-227</sub> loading control; lane 2 (**a**) and lane 1 (**d**) anti-PA serum control (no PA added).

Given the likelihood that ATR is conserved among different species, it is of interest to note that the product of the mouse homologue of ATR/TEM8 (GenBank accession number AK013005) is highly related to the human clones, sharing greater than 98% sequence identity within the reported extracellular domain (data not shown). Furthermore, consistent with the observation that the anthrax toxin receptor is found in a variety of cell lines, ATR and/or TEM8 is expressed in a number of different tissues including the central nervous system, heart, lung and lymphocytes (data not shown; and see NCBI UniGene cluster Hs.8966 at <http://www.ncbi.nlm.nih.gov/UniGene/clust.cgi?ORG=Hs&CID=8966&OPT=text>).

To confirm that PA binds directly to ATR, co-immunoprecipitations were performed with an extracellular fragment of ATR and either the wild-type or a mutant form of PA deficient in receptor binding. A fusion protein consisting of a hexahistidine tag, a T7 tag, and amino acids 41–227 of ATR (the I domain) was expressed and purified from *Escherichia coli* cells. When mixed with wild-type PA, this construct, T7–ATR<sub>41–227</sub>, was precipitated with polyclonal anti-PA serum (Fig. 3a, lane 3). The interaction between PA and T7–ATR<sub>41–227</sub> was impaired by the presence of EDTA (Fig. 3a, lane 5), demonstrating that the interaction requires divalent cations and suggesting that the MIDAS motif of ATR is critical for binding PA. In addition, a fusion protein consisting of glutathione S-transferase (GST) and the receptor-binding domain 4 (D4)<sup>25,26</sup> of PA (GST–D4) bound T7–ATR<sub>41–227</sub>, whereas GST did not (Fig. 3b). PA–N682S, a mutant form of PA with residue Asn 682 replaced with Ser, is impaired in its ability to bind and intoxicate cells (Fig. 3c and d), and was unable to bind to T7–ATR<sub>41–227</sub> (Fig. 3a, lane 4). These experiments demonstrate a direct and specific interaction between the VWA/I domain of ATR and the receptor-binding domain of PA.

Given this direct interaction, we reasoned that ATR<sub>41–227</sub> might protect CHO-K1 cells from being killed by PA and LF<sub>N</sub>–DTA. We tested this idea by mixing cells with an increasing amount of T7–ATR<sub>41–227</sub> in the presence of a constant amount of PA and LF<sub>N</sub>–DTA, and then measuring the subsequent effect on protein synthesis (Fig. 4). T7–ATR<sub>41–227</sub> was an effective inhibitor of toxin action, inhibiting toxin activity by 50% and 100% at concentrations of 80 and 500 nM, respectively. T7–ATR<sub>41–227</sub> did not, however, inhibit diphtheria toxin (data not shown).

We have identified a 368-amino-acid human protein, ATR, that contains a single extracellular VWA/I domain and serves as the cellular receptor for anthrax toxin. The clone that encodes this

protein may be one of several alternatively spliced messenger RNA transcripts, including TEM8, that result from a primary mRNA. Because TEM8 also contains the extracellular VWA/I domain, which binds directly to PA, we predict that this clone may also function as a PA receptor. The identification of ATR now allows for a more detailed investigation of the mechanism of uptake by cells of anthrax toxin. Furthermore, that the soluble VWA/I domain of ATR inhibits toxin action, coupled with the use of the cloned receptor as a tool for identifying inhibitors of the PA–receptor interaction, holds promise for the development of new approaches for the treatment of anthrax.

**Note added in proof:** A further apparently full-length ATR/TEM8-related cDNA clone has been reported (GenBank accession code BC012074), which encodes a protein with yet another C-terminal end. □

## Methods

### Mutagenesis and characterization of CHO-K1 cells

About  $5 \times 10^7$  CHO-K1 cells were treated at 37 °C for 7 h with medium containing  $10 \mu\text{g ml}^{-1}$  ICR-191 (Sigma) then washed twice. After 4 d, surviving cells were replated and incubated for 3 d with medium containing  $8 \mu\text{g ml}^{-1}$  PA and  $10 \text{ ng ml}^{-1}$  LF<sub>N</sub>–DTA. Surviving single-cell clones were isolated 14 d later. CHO-R1.1 cells were assayed for their sensitivity to DT intoxication by measuring incorporation of  $^3\text{H}$ -leucine into cellular proteins after exposure to the toxin<sup>16</sup>. Flow cytometry analysis was performed after incubating cells at 4 °C for 2 h in medium containing 40–80 nM OGPA (PA–K563C coupled to Oregon Green maleimide (Molecular Probes)). The cells were then washed twice with medium and analysed with a Becton Dickinson FACSCalibur flow cytometer.

### cDNA complementation

About  $5 \times 10^5$  CHO-R1.1 cells were transduced with  $\sim 10^7$  infectious units (complexity of library =  $2 \times 10^6$  independent clones) of the pLIB-based cDNA library (Clontech) produced in the 293GPG packaging cell line<sup>27</sup>. Three days later, cells were incubated with medium containing 80 nM OGPA and the top 0.1% of fluorescent cells were then isolated by sorting using a Becton Dickinson FACS Vantage SE instrument. These cells were expanded and subjected to four additional rounds of sorting using OGPA as above, as well as a 1:500 dilution of a rabbit anti-PA polyclonal serum along with a 1:500 dilution of an APC-conjugated secondary antibody (Molecular Probes). OGPA single positive (round 2) or OGPA/APC double positive (rounds 3–5) cells were recovered (the top 20%, 1%, 5% and 50% of fluorescent cells for rounds 2, 3, 4 and 5, respectively) and expanded after each round of sorting. The cDNA inserts contained within these cells were recovered by PCR amplification of genomic DNA samples with oligonucleotide primers specific for the MLV vector according to the manufacturer's instructions (Clontech). Each cDNA was sub-cloned between the *NotI* and *Sall* restriction enzyme sites of pLIB and the resulting plasmids were cotransfected into 293 cells with MLV gag/pol and VSV-G expression plasmids pMD.ori.gagpol and pMD.G<sup>28</sup>. Resulting pseudotyped virus particles were used to infect CHO-R1.1 and CHO-K1 cells followed by OGPA staining and FACS analysis as above. All sequencing was performed by the McArdle Laboratory macromolecular core.

### Cloning and expression of T7–ATR<sub>41–227</sub>

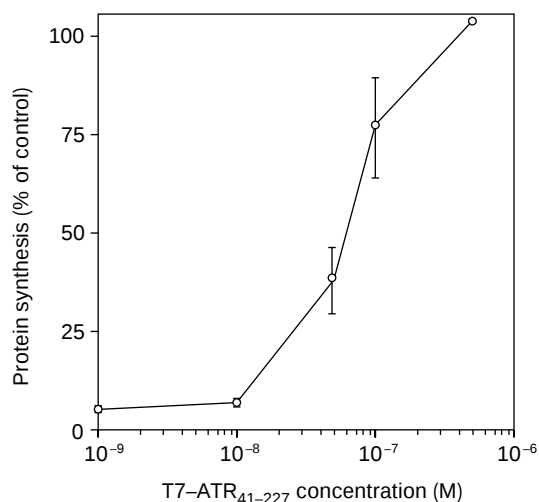
A DNA fragment encoding amino acids 41–227 of ATR was cloned into the *Bam*HI and *Eco*RI sites of pET28A (Novagen) to generate pET28A–ATR<sub>41–227</sub>. BL21 (DE3) cells (Stratagene) containing pET28A–ATR<sub>41–227</sub> were grown at 37 °C to an absorbance at 600 nm ( $A_{600}$ ) of 0.6, induced with 1 mM isopropyl- $\beta$ -D-thiogalactopyranoside for 4 h and collected by centrifugation. The cells from 1.5 l of culture were resuspended in 25 ml of 50 mM Tris–HCl buffer at pH 8.0, 2 mM dithiothreitol (DTT) and 1 mM phenylmethylsulphonyl fluoride, and were passed through a French press. We added 1 mg of DNase I (Roche) to the cell lysate, which we then sonicated for 1 min and centrifuged at 21,000g for 20 min. The pellet was resuspended in 25 ml of 50 mM Tris–HCl at pH 8.0 and 2 mM DTT, and centrifuged at 21,000g for 20 min. This wash step was repeated once. T7–ATR<sub>41–227</sub> was solubilized and folded essentially as described previously<sup>29</sup>.

### Isolation of PA–N682S

The DNA encoding domain 4 of PA was mutagenized by error-prone PCR<sup>31</sup>. Clones were expressed in *E. coli*, and lysates derived from these clones were added to CHO-K1 cells in combination with LF<sub>N</sub>–DTA. Clones corresponding to lysates that did not kill CHO-K1 cells were sequenced and the N682S mutant clone was further characterized here.

### Co-immunoprecipitation of PA and T7–ATR<sub>41–227</sub>

A mixture of 5  $\mu\text{g}$  PA (wild type or N682S) and 2  $\mu\text{g}$  T7–ATR<sub>41–227</sub> (in 20 mM Tris–HCl at pH 8.0, 150 mM NaCl and 0.1 mg ml<sup>−1</sup> bovine serum albumin) was incubated on ice for 30 min in the presence or absence of 2 mM EDTA. Anti-PA polyclonal serum (10  $\mu\text{l}$ ) was added to this solution and incubated on ice for an additional 1 h. Protein A agarose (Santa Cruz Biotechnology) was added and the solution was rotated at 4 °C for 1 h, then washed four times with 20 mM Tris–HCl at pH 8.0 and 150 mM NaCl. About one-third of the mixture was subjected to SDS–PAGE, transferred to nitrocellulose and probed with anti-T7 antibody conjugated to horseradish peroxidase (Novagen).



**Figure 4** T7–ATR<sub>41–227</sub> protects cells from killing by PA with LF<sub>N</sub>–DTA. CHO-K1 cells were incubated at 37 °C for 4 h with  $10^{-10}$  M PA,  $2.5 \times 10^{-11}$  M LF<sub>N</sub>–DTA and increasing amounts of T7–ATR<sub>41–227</sub>. Toxin sensitivity was determined by measuring the inhibition of protein synthesis as determined in Fig. 1c.



# GST–D4 pull-down assay

DNA encoding amino acids 595–735 of PA (domain 4) was cloned into pGEX-4T-1 (Pharmacia Biotechnology). GST–D4 was coupled to glutathione sepharose at 4 mg ml<sup>–1</sup> GST–D4 according to the manufacturer's instructions (Pharmacia Biotechnology). GST or GST–D4 coupled to glutathione sepharose was mixed with 2 µg of T7–ATR<sub>41–227</sub> and 250 µg of *E. coli* extract in 250 µl for 1 h at 4 °C. The beads were washed four times with 20 mM Tris–HCl at pH 8.0 and 150 mM NaCl. One-half of the suspension was subjected to SDS–PAGE, transferred to nitrocellulose, and probed with anti-T7 antibody coupled to horseradish peroxidase.

Received 21 September; accepted 10 October 2001.

Published online 23 October 2001, DOI: 10.1038/n35101999

1. Leppla, S. H. Anthrax toxin edema factor: a bacterial adenylate cyclase that increases cAMP concentrations in eukaryotic cells. *Proc. Natl Acad. Sci. USA* **79**, 3162–3166 (1982).
2. O'Brien, J. *et al.* Effects of anthrax toxin components on human neutrophils. *Infect. Immun.* **47**, 306–310 (1985).
3. Duesbery, N. S. *et al.* Proteolytic inactivation of MAP-kinase-kinase by anthrax lethal factor. *Science* **280**, 734–737 (1998).
4. Pellizzari, R. *et al.* Anthrax lethal factor cleaves MKK3 in macrophages and inhibits the LPS/IFN-γ-induced release of NO and TNFα. *FEBS Lett.* **462**, 199–204 (1999).
5. Friedlander, A. M. Macrophages are sensitive to anthrax lethal toxin through an acid-dependent process. *J. Biol. Chem.* **261**, 7123–7126 (1986).
6. Molloy, S. S. *et al.* Human furin is a calcium-dependent serine endoprotease that recognizes the sequence Arg–X–X–Arg and efficiently cleaves anthrax toxin protective antigen. *J. Biol. Chem.* **267**, 16396–16402 (1992).
7. Milne, J. C. *et al.* Anthrax protective antigen forms oligomers during intoxication of mammalian cells. *J. Biol. Chem.* **269**, 20607–20612 (1994).
8. Elliott, J. L., Mogridge, J. & Collier, R. J. A quantitative study of the interactions of *Bacillus anthracis* edema factor and lethal factor with activated protective antigen. *Biochemistry* **39**, 6706–6713 (2000).
9. Gordon, V. M., Leppla, S. H. & Hewlitt, E. L. Inhibitors of receptor-mediated endocytosis block the entry of *Bacillus anthracis* adenylate cyclase toxin but not that of *Bordetella pertussis* adenylate cyclase toxin. *Infect. Immun.* **56**, 1066–1069 (1988).
10. Blaustein, R. O., Koehler, T. M., Collier, R. J. & Finkelstein, A. Anthrax toxin: Channel-forming activity of protective antigen in planar phospholipid bilayers. *Proc. Natl Acad. Sci. USA* **86**, 2209–2213 (1989).
11. Koehler, T. M. & Collier, R. J. Anthrax toxin protective antigen: low-pH-induced hydrophobicity and channel formation in liposomes. *Mol. Microbiol.* **5**, 1501–1506 (1991).
12. Milne, J. C. & Collier, R. J. pH-dependent permeabilization of the plasma membrane of mammalian cells by anthrax protective antigen. *Mol. Microbiol.* **10**, 647–653 (1993).
13. Escuyer, V. & Collier, R. J. Anthrax protective antigen interacts with a specific receptor on the surface of CHO-K1 cells. *Infect. Immun.* **59**, 3381–3386 (1991).
14. Friedlander, A. M. *et al.* Characterization of macrophage sensitivity and resistance to anthrax lethal toxin. *Infect. Immun.* **61**, 245–252 (1993).
15. Taft, S. A., Liber, H. L. & Skopek, T. R. Mutational spectrum of ICR-191 at the *hprt* locus in human lymphoblastoid cells. *Environ. Mol. Mutagen.* **23**, 96–100 (1994).
16. Milne, J. C., Blanke, S. R., Hanna, P. C. & Collier, R. J. Protective antigen-binding domain of anthrax lethal factor mediates translocation of a heterologous protein fused to its amino- or carboxy-terminus. *Mol. Microbiol.* **15**, 661–666 (1995).
17. Burns, J. C. *et al.* Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc. Natl Acad. Sci. USA* **90**, 8033–8037 (1993).
18. Kitamura, T. *et al.* Efficient screening of retroviral cDNA expression libraries. *Proc. Natl Acad. Sci. USA* **92**, 9146–9150 (1995).
19. Whitehead, I., Kirk, H. & Kay, R. Expression cloning of oncogenes by retroviral transfer of cDNA libraries. *Mol. Cell. Biol.* **15**, 704–710 (1995).
20. St Croix, B. *et al.* Genes expressed in human tumor endothelium. *Science* **289**, 1197–1202 (2000).
21. Dickeson, S. K. & Santaro, S. A. Ligand recognition by the I domain-containing integrins. *Cell Mol. Life Sci.* **54**, 556–566 (1998).
22. Lee, J. O., Rieu, P., Arnaout, M. A. & Liddington, R. Crystal structure of the a domain from the subunit of integrin CR3 (CD11b/CD18). *Cell* **80**, 631–638 (1995).
23. Molloy, S. S., Anderson, E. D., Jean, F. & Thomas, G. Bi-cycling the furin pathway: from TGN localization to pathogen activation and embryogenesis. *Trends Cell Biol.* **9**, 28–35 (1999).
24. Beauregard, K. E., Wimer-Mackin, S., Collier, R. J. & Lencer, W. I. Anthrax toxin entry into polarized epithelial cells. *Infect. Immun.* **67**, 3026–3030 (1999).
25. Petosa, C. *et al.* Crystal structure of the anthrax toxin protective antigen. *Nature* **385**, 833–838 (1997).
26. Varughese, M., Teixeira, A. V., Liu, S. & Leppla, S. H. Identification of a receptor-binding region within domain 4 of the protective antigen component of anthrax toxin. *Infect. Immun.* **67**, 1860–1865 (1999).
27. Ory, D. S., Neugeboren, B. A. & Mulligan, R. C. A stable human-derived packaging cell line for production of high titer retrovirus/vesicular stomatitis virus G pseudotypes. *Proc. Natl Acad. Sci. USA* **93**, 11400–11406 (1996).
28. Snitkovsky, S. *et al.* A TVA-single-chain antibody fusion protein mediates specific targeting of a subgroup A avian leukosis virus vector to cells expressing a tumor-specific form of epidermal growth factor receptor. *J. Virol.* **74**, 9540–9545 (2000).
29. Legge, G. B. *et al.* NMR solution structure of the inserted domain of human leukocyte function associated antigen-1. *J. Mol. Biol.* **295**, 1251–1264 (2000).
30. Emsley, J., King, S. L., Bergelson, J. M. & Liddington, R. C. Crystal structure of the I domain from Integrin α2β1. *J. Biol. Chem.* **272**, 28512–28517 (1997).
31. Mogridge, J., Mourez, M. & Collier, R. J. Involvement of domain 3 in oligomerization by the protective antigen moiety of anthrax toxin. *J. Bacteriol.* **183**, 2111–2116 (2001).

## Acknowledgements

We thank K. Schell and members of the flow cytometry facility at the University of Wisconsin Comprehensive Cancer Center for performing the cell sorting. This work was

supported by a grant to J.A.T.Y. and R.J.C. from the National Institutes of Health. J.M. was supported in part by a Medical Research Council (Canada) postdoctoral fellowship.

## Competing interests statement

The authors declare competing financial interests: details accompany the paper on Nature's website (<http://www.nature.com>).

Correspondence and requests for materials should be addressed to J.A.T.Y. (e-mail: [young@oncology.wisc.edu](mailto:young@oncology.wisc.edu)). The ATR cDNA clone has been deposited in GenBank (accession code AF421380).

# Crystal structure of the anthrax lethal factor

Andrew D. Pannifer\*†, Thiang Yian Wong‡, Robert Schwarzenbacher‡, Martin Renatus‡, Carlo Petosa\*†, Jadwiga Bienkowska†§, D. Borden Lacy||, R. John Collier||, Sukjoon Park¶, Stephen H. Leppla¶, Philip Hanna# & Robert C. Liddington‡

\* Biochemistry Department, University of Leicester, Leicester LE1 7RH, UK

‡ The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, California 92037, USA

§ Dana Farber Cancer Institute, Boston, Massachusetts 02115, USA

|| Department of Microbiology and Molecular Genetics, Harvard Medical School, 200 Longwood Avenue, Boston, Massachusetts 02115, USA

¶ National Institute of Dental and Craniofacial Research, National Institutes of Health, 9000 Rockville Pike, Bethesda, Maryland 20892, USA

# Department of Microbiology and Immunology, University of Michigan Medical School, 5641 Medical Science II, Ann Arbor, Michigan 48109, USA

Lethal factor (LF) is a protein (relative molecular mass 90,000) that is critical in the pathogenesis of anthrax<sup>1–3</sup>. It is a highly specific protease that cleaves members of the mitogen-activated protein kinase kinase (MAPKK) family near to their amino termini, leading to the inhibition of one or more signalling pathways<sup>4–6</sup>. Here we describe the crystal structure of LF and its complex with the N terminus of MAPKK-2. LF comprises four domains: domain I binds the membrane-translocating component of anthrax toxin, the protective antigen (PA); domains II, III and IV together create a long deep groove that holds the 16-residue N-terminal tail of MAPKK-2 before cleavage. Domain II resembles the ADP-ribosylating toxin from *Bacillus cereus*, but the active site has been mutated and recruited to augment substrate recognition. Domain III is inserted into domain II, and seems to have arisen from a repeated duplication of a structural element of domain II. Domain IV is distantly related to the zinc metalloprotease family, and contains the catalytic centre; it also resembles domain I. The structure thus reveals a protein that has evolved through a process of gene duplication, mutation and fusion, into an enzyme with high and unusual specificity.

Anthrax poses a significant threat as an agent of biological warfare and terrorism<sup>7</sup>. Inhalational anthrax, in which spores of *Bacillus anthracis* are inhaled, is almost always fatal, as diagnosis is rarely possible before the disease has progressed to a point where antibiotic treatment is ineffective. The major virulence factors of *B. anthracis* are a poly-D-glutamic acid capsule and anthrax toxin<sup>1</sup>. Anthrax toxin consists of three distinct proteins that act in concert: two enzymes, LF and oedema factor (an adenylate

† Present addresses: AstraZeneca, Alderley Park, Macclesfield SK10 4TG, UK (A.P.); EMBL Grenoble Outstation, 38042 Grenoble Cedex 9, France (C.P.); Boston University, 36 Cummington Street, Boston, Massachusetts 02215, USA (J.B.).

cyclase); and PA. The PA is a four-domain protein<sup>8</sup> that binds a host cell-surface receptor by its carboxy-terminal domain; cleavage of its N-terminal domain by a furin-like protease allows PA to form heptamers that bind the toxic enzymes with high affinity through homologous N-terminal domains. The complex is endocytosed; acidification of the endosome leads to membrane insertion of the PA heptamer by forming a 14-stranded  $\beta$ -barrel<sup>8,9</sup>, followed by translocation of the toxic enzymes into the cytosol by an unknown mechanism. The binary combination of PA and LF ('lethal toxin') is sufficient to induce rapid death in animals when given intravenously, and certain metalloprotease inhibitors block the effects of the toxin *in vitro*<sup>10</sup>. Thus, LF is a potential target for therapeutic agents that would inhibit its catalytic activity or block its association with PA.

The MAPKK family of proteins are the only known cellular substrates of LF<sup>4–6</sup>. Cleavage by LF near to their N termini removes the docking sequence for the downstream cognate MAP kinase. The effect of lethal toxin on tumour cells, for example, is to inhibit tumour growth and angiogenesis, most probably by inhibiting the MAPKK-1 and MAPKK-2 pathways<sup>11</sup>. However, the primary cell type affected in anthrax pathogenesis is the macrophage<sup>12,13</sup>. Here, recent evidence shows that low levels of lethal toxin cleave MAPKK-3, inhibiting release, but not production, of the pro-inflammatory mediators, NO and tumour necrosis factor- $\alpha$  (TNF- $\alpha$ )<sup>14</sup>. In contrast, high levels of lethal toxin lead to lysis of macrophages within a few hours, by an unknown mechanism<sup>12</sup>. These observations suggest that at an early stage in infection, lethal toxin may reduce (or delay) the immune response, whereas at a late stage in infection, high titres of the bacterium in the bloodstream trigger macrophage lysis and the sudden release of high levels of NO and TNF- $\alpha$ . This may explain the symptoms before death, which resemble those of septic shock<sup>2</sup>.

We solved the structure of LF in two crystal forms with very different packing environments. Nevertheless, differences in tertiary and quaternary structure are small, and the better diffracting crystal form has been refined to 2.2 Å resolution (Figs 1 and 2). The molecule is 100 Å tall and 70 Å wide at its base, with domain I perched on top of the other three domains, which are intimately connected and probably comprise a single folding unit. The only contacts between domain I and the rest of the molecule are with domain II, and these chiefly involve charged polar and water-mediated interactions. The nature of the interface is consistent with the ability of a recombinant N-terminal fragment (residues 1–254) to be expressed as a soluble folded domain that maintains the ability to bind PA and enables the translocation of heterologous fusion proteins into the cytosol<sup>15,16</sup>.

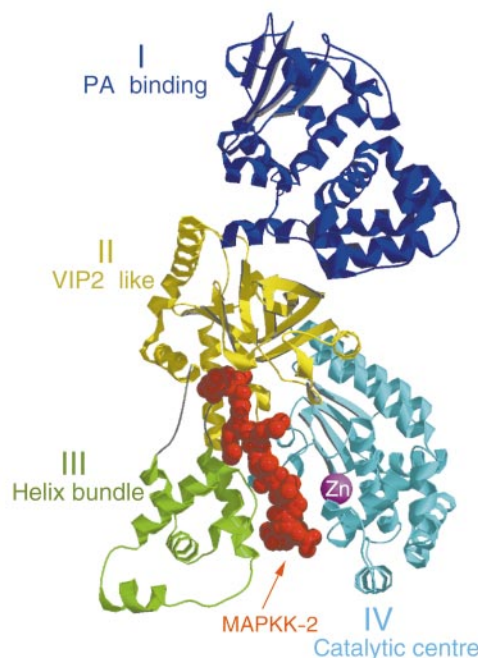
Domain I consists of a 12-helix bundle that packs against one face of a mixed four-stranded  $\beta$ -sheet, with a large (30-residue) ordered loop, L1, between the second and third  $\beta$ -strands forming a flap over the distal face of the sheet. Although there is no detectable sequence homology between domains I and IV, the sheet and two abutting helices can be superposed with a root mean square difference (RMSD) of 2.6 Å. The hallmark metalloprotease motif HExxH is not conserved in domain I; it is replaced by the sequence YEIGK, so Zn<sup>2+</sup> coordination is not possible. The docking site on domain I for PA is unknown, but the integrity of the folded domain seems to be required, because a series of insertion and point mutants of buried residues in domain I (that presumably disrupt the fold) abrogate binding of PA and toxicity<sup>17,18</sup>.

An abrupt turn at the end of the last helix of domain I leads directly into the first helix of domain II (residues 263–297 and 385–550). Although sequence-based comparisons failed to yield any homology, the structural similarity with the catalytic domain of the *B. cereus* toxin, VIP2 (Protein Data Bank accession code 1QS2), is outstanding. Domain II and VIP2 superpose with an RMSD of 3.3 Å and a sequence identity of 15%, as determined by DALI<sup>19</sup>. VIP2 contains an NAD-binding pocket and conserved residues involved

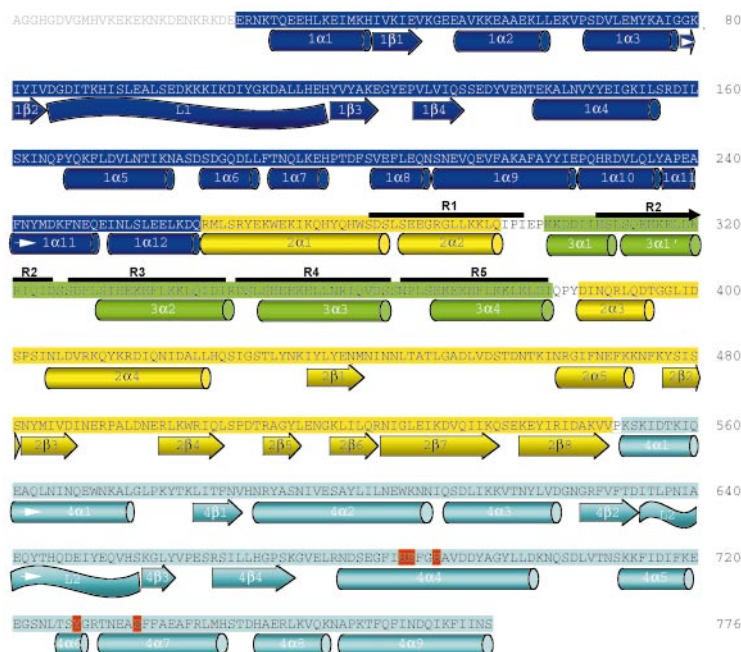
in NAD binding and catalysis. Domain II lacks these conserved residues; moreover, a critical glutamic acid that is conserved throughout the family of ADP ribosylating toxins<sup>20</sup> is replaced by a lysine (K518). We therefore expect that domain II does not have ADP-ribosylating activity.

Domain III is a small  $\alpha$ -helical bundle with a hydrophobic core (residues 303–382), inserted at a turn between the second and third helices of domain II. Sequence analysis had revealed the presence of a 101-residue segment comprising five tandem repeats (residues 282–382), and suggested that repeats 2–5 arose from a duplication of repeat 1. The crystal structure reveals that repeat 1 actually forms the second helix-turn element of domain II, whereas repeats 2–5 form the four helix-turn elements of the helical bundle, suggesting a mechanism of creating a new protein domain by the repeated replication of a short segment of the parent domain. Domain III is required for LF activity, because insertion mutagenesis and point mutations of buried residues in this domain abrogate function<sup>17</sup>. It makes limited contact with domain II, but shares a hydrophobic surface with domain IV. Its location is such that it severely restricts access to the active site by potential substrates such as the loops of a globular protein; that is, it contributes towards specificity for a flexible 'tail' of a protein substrate. It also contributes sequence specificity by making specific interactions with the substrate (see below).

Domain IV (residues 552–776) consists of a nine-helix bundle packed against a four-stranded  $\beta$ -sheet. Sequence comparisons had failed to detect any homology with other proteins of known structure beyond the HExxH motif. Our three-dimensional structure reveals that the  $\beta$ -sheet and the first six helices can be superposed with those of the metalloprotease thermolysin, with an RMSD of 4.9 Å over 131 residues. Large insertions and deletions occur elsewhere within the loops connecting these elements, so that the overall shapes of the domains are quite different. In particular, a



**Figure 1** Stereo ribbon representation of LF, coloured by domain. The MAPKK-2 substrate is shown as a red ball-and-stick model, and the Zn<sup>2+</sup> ion is labelled. The RMSD (C $\alpha$ ) between the cubic and monoclinic crystal forms is 1.18 Å. The principal differences lie in the position of the helical bundle of domain IV, which undergoes a rigid body shift of ~2 Å relative to the  $\beta$ -sheet, and a smaller shift of the  $\beta$ -sheet of domain I. The third helical element of domain III is invisible in the monoclinic crystals, but can be seen in cubic crystals, albeit with high *B*-factors. It is possible that this mobile amphipathic helix has a role in the membrane-inserting properties of LF observed *in vitro*<sup>28</sup>. Figure prepared with MOLSCRIPT, RENDER and RASTER3D<sup>29–31</sup>.



**Figure 2** Sequence of LF, coloured as in Fig. 1, with secondary structure indicated. The N-terminal 26 residues are invisible in our electron density maps, and we presume that they are disordered. All other residues are ordered in at least one crystal form. The five sequence repeats are indicated R1–R5. The zinc-coordinating residues are highlighted in

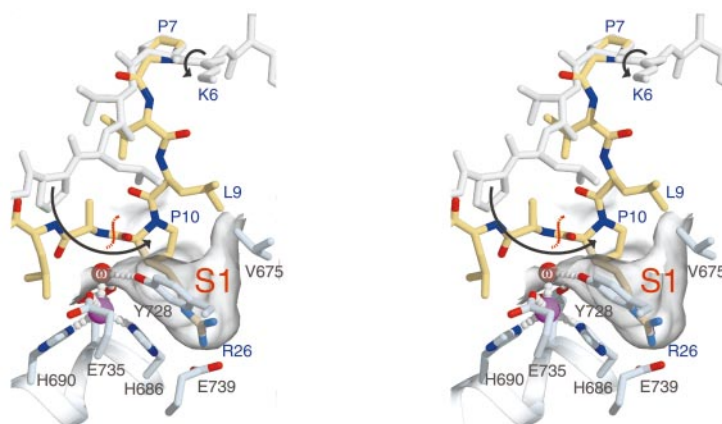
red. The sequence of domain I is homologous to the PA-binding domain of oedema factor, and its three-dimensional structure will be similar. For domain II, secondary structure nomenclature matches that of the N-terminal domain of VIP2.

large ordered loop (L2) inserted between strands 4β2 and 4β3 of the sheet partly obscures the active site, packs against domain II, and provides a buttress for domain III.

A zinc ion ( $\text{Zn}^{2+}$ ) is coordinated tetrahedrally by a water molecule and three protein side chains (Fig. 3), in an arrangement typical of the thermolysin family. Two coordinating residues are the histidines from the HExxH motif (His 686 and His 690) lying on one helix (4α4), as expected. Our structure reveals that the third coordinating residue is Glu 735 from helix 4α6. Glu 687 from the HExxH motif lies 3.5 Å from the water molecule, well positioned to act as a general base to activate the zinc-bound water during catalysis. The hydroxyl group of a tyrosine residue (Tyr 728) forms a strong hydrogen bond (O–O distance 2.6 Å) to the water molecule, on the opposite side of

Glu 687, and probably functions as a general acid to protonate the amine leaving group.

A broad deep groove, 40 Å long, is contiguous with the active site centre; it is created by the vestigial NAD-binding pocket of domain II and by the interface between domains II, III and IV. The groove has in general a negative potential, containing clusters of glutamic acid/aspartic acid, as well as glutamine/asparagine residues. We introduced peptides corresponding to the N-terminal 16 residues of MAPKK-2 into crystals of LF. In the absence of  $\text{Zn}^{2+}$ , we observed continuous electron density that can accommodate the entire 16-residue peptide, extending from domain II through the catalytic centre and ending at the distal end of domain IV (Fig. 4). We observed similar peptide density in the presence of  $\text{Zn}^{2+}$  in peptide-



**Figure 3** Stereo representation of the active site centre of LF. Zinc-coordinating residues are shown in ball-and-stick representation with the  $\text{Zn}^{2+}$  ion in magenta and the zinc-bound water labelled  $\omega$ . Zinc–ligand bond distances are  $2.1 \pm 0.1$  Å (mean  $\pm$  s.d.). The molecular surface defining the S1 substrate-binding pocket is indicated. Part of the MAPKK-2 model derived crystallographically is shown in grey. A hypothetical model of the productive complex is shown in yellow, blue and red. A rotation about the Lys 6–Pro 7 main chain

would bring Pro 10 into the S1 pocket, positioning the scissile bond close to the zinc-bound water, as indicated. The equivalent residue to Pro 10 in MAPKK-3 is Arg 26, which has a side chain that is also modelled in the S1 pocket, where it could form a salt bridge with the buried Glu 739. A hydrophobic residue (Leu 9 in MAPKK-2) is favoured at the P1' position, and our model suggests that it can interact with Val 675, at the top of the S1 pocket. Figs 3 and 4 were prepared with SPOCK<sup>32</sup>.



soaked crystals of a mutant (E687C) that lacks catalytic activity but that binds  $\text{Zn}^{2+}$  normally<sup>10</sup> (data not shown). A brief (5 min) soak of the wild-type complex in  $\text{Zn}^{2+}$  led to the disappearance of most of the peptide electron density and a strong peak at the  $\text{Zn}^{2+}$  position, consistent with activation of the proteolytic machinery. We built a peptide model that shows convincing agreement with the electron density lying in the direction expected for a thermolysin family member. Residues 1–3 sit wholly within the groove of domain II, surrounded by a cluster of asparagine and glutamine residues. Residues 4–10 are sandwiched between domains III and IV. A strongly acidic region of the protein surface rationalizes the preference for basic residues in the MAPKK substrates located 3–6 residues upstream of the cleavage site<sup>6</sup>. Arg 4 makes specific hydrogen bonds with domain III, and Pro 7 points towards an acidic cavity large enough to accommodate the lysine or arginine side chains found in the equivalent positions in MAPKK-1 and -3. Beyond Pro 7, the peptide makes fewer specific contacts.

This is, to our knowledge, the first structural example of a protease in complex with its uncleaved natural substrate. The closest main-chain approach to the  $\text{Zn}^{2+}$  ion is the scissile bond following Pro 10. However, it is about 6 Å from the zinc-bound water, suggesting that we have captured a 'pre-cleavage' complex. Such complexes have been predicted on the basis of kinetic data<sup>21</sup> but have not previously been observed crystallographically, presumably because, for the shorter recognition sequences of typical proteases, this binding mode is much weaker. Modelling studies show that a rotation about a main-chain dihedral angle at Lys 6 would allow Pro 10 to swing down into the S1 specificity pocket to generate the productive cleavage complex (Fig. 3). This conformational change would not disturb the specific interactions between LF and the N-terminal half of the MAPKK tail. The productive conformation requires an abrupt 90° turn in the peptide chain (favoured by a proline residue) at the P1 position, which would explain why a proline-to-alanine point mutant is not cleaved efficiently<sup>4</sup>. The S1 pocket is very pronounced—hydrophobic at its neck and acidic at its base—appropriate to fit the arginine or lysine side chains found at the P1 position in other MAPKK members<sup>6</sup>. Although the length of the pocket fits the N termini of MAPKK-1 and -2 very well, the pocket is not closed at its N-terminal end, so longer tails, such as those of other MAPKK members, could simply protrude beyond the pocket. At its C terminus, where the globular domain of MAPKK

begins, the peptide just protrudes from between domains III and IV. These surfaces may thus provide additional docking sites for other domains on the MAPKK, because fragments of MAPKK-1 lacking the N-terminal tail also bind to LF in a two-hybrid screen<sup>5</sup>.

The atomic resolution structure of LF described here demonstrates that the remarkable specificity for MAPKK family members arises in large part from the existence of an extended substrate-binding pocket created by gene duplication and fusion, which recognizes both the structural and sequence properties of its substrate. This information can be used in the design of therapeutic agents that would block the activity of LF *in vivo*. Our structure may also allow the design of mutant LF molecules with altered specificities for different members of the MAPKK family and hence improved anti-tumour properties, for example. □

## Methods

### Structure of the monoclinic crystal form

LF was purified as described elsewhere<sup>22</sup>. A mixture of cubic and monoclinic crystals were grown from 13 mg ml<sup>-1</sup> LF, 1.7–1.9 M ammonium sulphate, 0.2 M Tris buffer at pH 7.5–8.0, by the hanging-drop vapour diffusion method. The predominant crystal form is cubic ( $I4_132$ ,  $a = 331$  Å); these crystals diffract to 3.8 Å resolution. Monoclinic crystals ( $P2_1$ ,  $a = 96.2$  Å,  $b = 137.5$  Å,  $c = 98.6$  Å,  $\beta = 98.4^\circ$ ) were selectively obtained by macroseeding, and these diffract to 2.2 Å resolution. Derivatives of the monoclinic form were prepared by transferring the crystals to a stabilization buffer containing 2 M lithium sulphate, 200 mM Tris at pH 7.5, and either 50 mM trimethyl-lead acetate for 8 h or 15 mM potassium hexaiodoplatinate for 45 min. Native data were collected on station 9.6 at the Synchrotron Radiation Source (SRS) (Daresbury, UK); a platinum derivative data set was collected on an R-Axis IV image plate, and a multiple (three)-wavelength anomalous diffraction (MAD) data set from the lead derivative was collected on BM14 at the European Synchrotron Radiation Facility (ESRF) (Grenoble). All data were collected at 100 K after transfer to a cryoprotectant containing 25% glycerol. Data were indexed, integrated and scaled with the HKL package<sup>23</sup> and heavy-atom sites were identified and refined with FHSICAL, FFT and MLPHARE from the CCP4 suite<sup>24</sup> or with the program SOLVE<sup>25</sup>. The final figure of merit was 0.51 at 3.2 Å. The rotational component of a two-fold noncrystallographic symmetry (NCS) axis was identified with POLARREN<sup>24</sup> and the translation with GETAX<sup>24</sup>. Model building was carried out by the graphics program TURBO-FRODO<sup>26</sup> and refined with CNS<sup>27</sup> using all data and a bulk solvent correction. A monomer mask was identified in a map solvent flattened with DM, allowing strict NCS constraints to be used in model building and map calculation. The *R*-factor converged at around 30%; NCS constraints were released and a further cycle of refinement resulted in a drop in the *R*<sub>free</sub> of almost 3%. Multiple cycles of refinement, rebuilding and addition of water molecules reduced the *R*-factors to final values of 0.225 (*R*<sub>work</sub>) and 0.26 (*R*<sub>free</sub>) at 2.2 Å. The resulting model fell within, or exceeded, the limits of all the quality criteria of the program PROCHECK<sup>24</sup>. The final model contains 1,497 residues (residues 27–773 of molecule A and 27–776 of molecule B), 934 water molecules, 2 sulphate ions and 2 zinc ions. A table of crystallographic statistics is provided in Supplementary Information.

### Structure of the cubic crystal form

We solved cubic structure by molecular replacement with MOLREP<sup>24</sup>, using the refined model of the  $P2_1$  crystal form converted to polyalanine as a search probe. The cubic crystals contained one molecule in the asymmetric unit and 86% solvent. The initial solution (*R*<sub>free</sub> = 0.49) underwent preliminary refinement using rigid body minimization and simulated annealing procedures in CNS<sup>27</sup> to *R*<sub>free</sub> = 0.38. Owing to the limited resolution (3.8 Å) and the rapid fall-off in diffraction intensity, the data were sharpened with an overall *B*-factor of 50 Å<sup>2</sup> (on I), to increase the weight of the high-resolution terms. The resulting difference and omit maps used for rebuilding revealed higher resolution features, such as improved side-chain density.

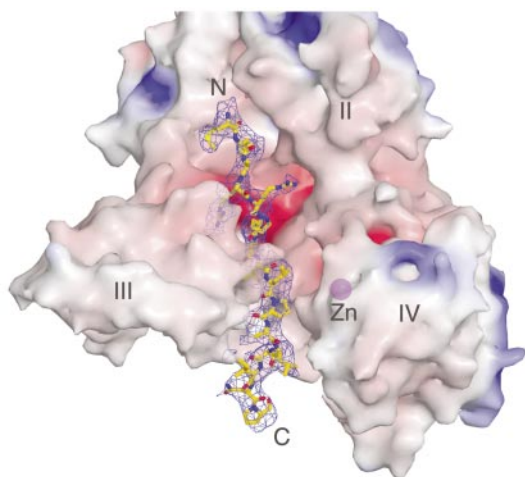
### Structure of peptide complex

The monoclinic crystals cracked on exposure to MAPKK peptides, so the cubic form was used for substrate binding studies. The highest peptide occupancy was achieved with a 5-min room-temperature soak of crystals in 10 mM MAPKK-2 peptide (MLARRKPVLP ALTINP) in cryoprotectant in the absence of zinc. A  $2F_o - F_c$  map revealed an additional elongated stretch of electron density in the active site crevice. A peptide model containing the MAPKK-2 sequence was built and subjected to further cycles of refinement and rebuilding with the programs REFMAC5<sup>24</sup> and TURBO-FRODO<sup>26</sup>. Periodically and for final maps, model bias was minimized by excluding the peptide from the phasing and carrying out a cycle of refinement ('simulated annealing omit map'<sup>27</sup>). Assuming an occupancy of 0.7, the peptide main chain *B*-factors refine to values similar to those of the protein. The final model for the cubic data has been refined to *R*<sub>free</sub> = 0.316 (0.402 in the 3.9–4.0 Å shell), *R*<sub>work</sub> = 0.295.

Received 16 July; accepted 18 September 2001.

Published online 23 October 2001, DOI: 10.1038/n35101998

1. Leppla, S. H. in *Comprehensive Sourcebook of Bacterial Protein Toxins* 2nd edn (eds Alouf, J. A. & Freer, J.) 243–263 (Academic, London, 1999).



**Figure 4** Molecular surface of LF coloured by charge (red, negative; blue, positive), with the model of the MAPKK-2 N-terminal peptide shown in ball-and-stick representation. The electron density surrounding the peptide is a 'simulated annealing omit map' (see Methods) calculated at 3.9 Å resolution and contoured at 0.7σ. Where the model or map would be hidden by the protein surface, the surface is rendered as translucent. A stereo version of this figure is provided in the Supplementary Information.

2. Dixon, T. C., Meselson, M., Guillemin, J. & Hanna, P. C. Medical progress: anthrax. *N. Engl. J. Med.* **341**, 815–862 (1999).
3. Pezard, C., Berche, P. & Mock, M. Contribution of individual toxin components to virulence of *Bacillus anthracis*. *Infect. Immun.* **59**, 3472–3477 (1991).
4. Duesbury, N. S. *et al.* Proteolytic inactivation of MAP-kinase-kinase by anthrax lethal factor. *Science* **280**, 734–737 (1998).
5. Vitale, G. *et al.* Anthrax lethal factor cleaves the N-terminus of MAPKKs and induces tyrosine/threonine phosphorylation of MAPKs in cultured macrophages. *Biochem. Biophys. Res. Comm.* **248**, 706–711 (1998).
6. Vitale, G., Bernardi, L., Napolitani, G., Mock, M. & Montecucco, C. Susceptibility of mitogen-activated protein kinase kinase family members to proteolysis by anthrax lethal factor. *Biochem. J.* **352**, 739–745 (2000).
7. Inglesby, T. V. *et al.* Anthrax as a biological weapon: medical and public health management. Working group on civilian biodefense. *J. Am. Med. Assoc.* **281**, 1735–1745 (1999).
8. Petosa, C., Collier, R. J., Klimpel, K. R., Leppla, S. H. & Liddington, R. C. Crystal structure of the anthrax toxin protective antigen. *Nature* **385**, 833–838 (1997).
9. Benson, E. L., Huynh, P. D., Finkelstein, A. & Collier, R. J. Identification of residues lining the anthrax protective antigen channel. *Biochemistry* **37**, 3941–3948 (1998).
10. Klimpel, K. R., Arora, N. & Leppla, S. H. Anthrax toxin lethal factor contains a zinc metalloprotease consensus sequence which is required for lethal toxin activity. *Mol. Microbiol.* **13**, 1093–1100 (1994).
11. Duesbery, N. S. *et al.* Suppression of ras-mediated transformation and inhibition of tumour growth and angiogenesis by anthrax lethal factor, a proteolytic inhibitor of multiple MEK pathways. *Proc. Natl Acad. Sci. USA* **98**, 4089–4094 (2001).
12. Friedlander, A. M. Macrophages are sensitive to anthrax lethal toxin through an acid-dependent process. *J. Biol. Chem.* **261**, 7123–7126 (1986).
13. Hanna, P. C., Acosta, D. & Collier, R. J. On the role of macrophages in anthrax. *Proc. Natl Acad. Sci. USA* **90**, 10198–10201 (1993).
14. Pellizzari, R., Guidi-Rontani, C., Vitale, G., Mock, M. & Montecucco, C. Anthrax lethal factor cleaves MKK3 in macrophages and inhibits the LPS/IFN $\gamma$ -induced release of NO and TNF $\alpha$ . *FEBS Lett.* **462**, 199–204 (1999).
15. Ballard, J. D., Collier, R. J. & Starnbach, M. N. Anthrax toxin-mediated delivery of a cytotoxic T-cell epitope *in vivo*. *Proc. Natl Acad. Sci. USA* **93**, 12531–12534 (1996).
16. Goletz, T. J. *et al.* Targeting HIV proteins to the major histocompatibility complex class I processing pathway with a novel gp120–anthrax toxin fusion protein. *Proc. Natl Acad. Sci. USA* **94**, 12059–12064 (1997).
17. Quinn, C. P., Singh, Y., Klimpel, K. R. & Leppla, S. H. Functional mapping of anthrax toxin lethal factor by in-frame insertion mutagenesis. *J. Biol. Chem.* **266**, 20124–20130 (1991).
18. Gupta, P., Singh, A., Chauhan, V. & Bhatnagar, R. Involvement of residues 147VYVEIGK153 in binding of lethal factor to protective antigen of *Bacillus anthracis*. *Biochem. Biophys. Res. Comm.* **280**, 158–163 (2001).
19. Holm, L. & Sander, C. Dali/FSSP classification of three-dimensional protein folds. *Nucleic Acids Res.* **25**, 231–234 (1997).
20. Carroll, S. F. & Collier, R. J. NAD binding site of diphtheria toxin: identification of a residue within the nicotinamide subsite by photochemical modification with NAD. *Proc. Natl Acad. Sci. USA* **81**, 3307–3311 (1984).
21. Mock, W. L. & Stanford, D. L. Arazoformyl dipeptide substrates for thermolysin. Confirmation of a reverse protonation catalytic mechanism. *Biochemistry* **35**, 7369–7377 (1996).
22. Park, S. & Leppla, S. H. Optimized production and purification of *Bacillus anthracis* lethal factor. *Protein Expr. Purif.* **18**, 293–302 (2000).
23. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
24. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
25. Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr. D* **55**, 849–861 (1999).
26. Roussel, A. & Cambillau, C. 86 (Silicon Graphics, Mountain View, California, 1991).
27. Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
28. Wang, X.-M., Mock, M., Ruysschaert, J.-M. & Cabiaux, V. Secondary structure of anthrax lethal toxin proteins and their interaction with large unilamellar vesicles (LUV): a Fourier-transform infrared spectroscopy approach. *Biochemistry* **35**, 14939–14946 (1996).
29. Bacon, D. J. & Anderson, W. F. A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graphics* **6**, 219–220 (1988).
30. Kraulis, P. J. Molscript—a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
31. Merrit, E. A. & Murphy, M. E. P. Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr. D* **50**, 869–873 (1994).
32. Christopher, J. A. *SPOCK: The Structural Properties Observation and Calculation Kit (Program Manual)* (The Center for Macromolecular Design, Texas A&M University, College Station, 1998).

**Supplementary information** is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

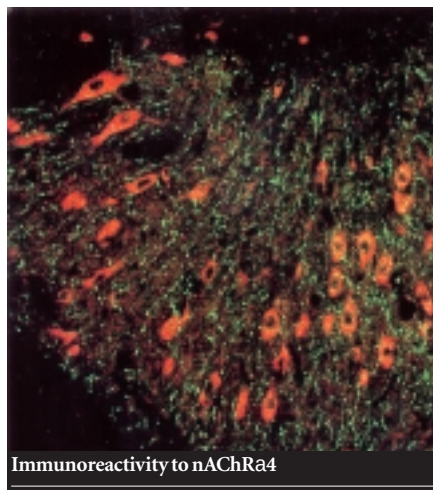
## Acknowledgements

We thank L. Bankston for discussion and D. Hsu for LF preparations. We thank and acknowledge the staff and facilities of the synchrotron sources at SSRL, Stanford; SRS, Daresbury; ESRF, Grenoble; APS, Chicago; and National Synchrotron Light Source, Brookhaven. Supported by grants from the Medical Research Council and the National Institutes of Health.

Correspondence and requests for materials should be addressed to R.C.L. (e-mail: [rlidding@burnham.org](mailto:rlidding@burnham.org)). Coordinates for the monoclinic LF crystal structure and the cubic LF–MAPKK2 complex have been deposited with the Protein Data Bank (accession codes 1J7N and 1JKY, respectively).

# Neuroscience on the brain

Antibodies, reagents and equipment for neuroscience research.



Immunoreactivity to nAChRα4

## nAChRα4 antibody

Chemicon [www.chemicon.com](http://www.chemicon.com)

*Antiserum suitable for double-labelling*

This antibody to nicotinic acetylcholine receptor  $\alpha_4$ R (No. AB5590) was raised in the guinea pig to a synthetic peptide of 20 amino acids corresponding to the intracellular domain of rat nAChR  $\alpha_4$ -subunit. Strong immunoreactivity to this antibody has been demonstrated in Purkinje cells in the cerebellum, the cerebral cortex, striatum, brainstem and spinal cord.

## Benchmark Digital

myNeuroLab.com [www.myneurolab.com](http://www.myneurolab.com)

*A stereotaxic instrument with digital scales*

The Benchmark Digital stereotaxic instrument features a linear position-sensing digital scales and an easy-to-read remote display centre. Designed with the same footprint and drive mechanisms as the standard line of Benchmark stereotaxic instruments, the new Digital unit is built for ease of use. Users simply move the probe to Bregma and touch the zeroing button on the display. Once all displays read zero, the probe is moved until each of the three axes dis-

plays the desired coordinates. A range of probe holders is available. In addition, configurations and accessories for mice, rats and other species are available. The unit is compatible with all stereotaxic instrument manufacturers.

## Ion channels brochure

Sigma-RBI [www.sigma-sial.com](http://www.sigma-sial.com)

*Reagents for ion channel research*

The new *Ion Channels* brochure lists over 200 products, including antibodies, under the headings potassium, sodium, chloride and water channels. The brochure is available from [www.sigma-aldrich.com/cellsignalling](http://www.sigma-aldrich.com/cellsignalling).

## Neuroscience antibodies/antigens

Biodesign International

[www.biodesign.com](http://www.biodesign.com)

*New reagents for Alzheimers' research*

The Biodesign neuroscience range has been broadened with the addition of 27 new antibodies and purified antigens. New introductions include polyclonals and antibodies for the study of amyloid precursor proteins, presenilins, neurofibrillary tangle and tau. For Parkinson's disease research new products include alpha-synuclein synphilin-1 and parkin.

## MBR microdialysis probes

Bioanalytical Systems Inc.

[www.bioanalytical.com](http://www.bioanalytical.com)

*Good as gold*

BAS have developed new brain microdialysis probes with intracerebral cannulae for use in transgenic mice, rats and other rodents. The light-weight MBR probes are secured in the guide by a low-insertion force elastomeric fit which accommodates the fragile skull of a mouse. MBR probes are available in 1 or 2 mm lengths, the membrane has a 38K molecular weight cutoff and is suitable for a broad range of neurotransmitters and drugs. Probe and guides are coated with 24-carat gold for compatibility with tissues and biological fluids.

## Nicastrin antibodies

Affinity BioReagents [www.bioreagents.com](http://www.bioreagents.com)

*A detection reagent for Alzheimer's research*

Nicastrin binds to the carboxyl terminus of the beta-amyloid precursor protein and helps to modulate production of amyloid beta peptide. ABR's antibody to nicastrin is optimized for western blots of human and mouse protein.

*These notes are compiled in the Nature office from information provided by the manufacturers.*

## ADVERTISEMENTS

**THE NOBEL PRIZE**  
The First 100 Years  
Nobel Prize  
edited by  
Agneta Wallin Levinovitz &  
Mika Ringertz  
(The Nobel Foundation)  
2001  
981-43-4884-1 95294 537  
981-43-4885-Xipho 95296 518  
World Scientific Imperial College Press  
[www.worldscientific.co.uk](http://www.worldscientific.co.uk) [www.icpress.co.uk](http://www.icpress.co.uk)

**DIVERSE SMALL MOLECULES READY FOR SCREENING**  
**10K-50K Compounds in Plates**  
• Next Day Delivery\*  
• Competitively Priced  
(\*Some Restrictions Apply)  
**Chembridge Corporation • 800.964.6143**  
858.451.7400 • 858.451.7401 fax • E-mail: [chem@chembridge.com](mailto:chem@chembridge.com)  
**www.chembridge.com**

## CUSTOMISED

immunoassays for  
research and production



**BioGenes**  
Gesellschaft für Biopolymere mbH

Internet [www.biogenes.de](http://www.biogenes.de)  
Berlin Germany  
Phone +49 (0)30-65 76 23 96

## Polyclonal Antibodies & Custom Peptides

Please contact us about complete services for your custom peptide and polyclonal antibody needs.

**ResGen™**  
Invitrogen Corporation

U.S. or Canada 800-533-4363  
FAX 256-536-9016

Check out our web site at  
[www.resgen.com](http://www.resgen.com)

## Proteasome inhibition?

~ epoxomicin ~  
~ PR39 ~ lactacystin ~  
~ clasto-lactacystin ~  $\beta$  ~  
~ lactone ~ P131 ~  
~ MG-132 ~ MG-262 ~

**www.proteasome.com**

AFFINITY Research Products Ltd.  
Mamhead Castle, Exeter, EX6 8HD, U.K.  
T.+44(0) 1626 891010 ~ F+44(0) 1626 891090  
~ Email [info@proteasome.com](mailto:info@proteasome.com) ~



## Contacts

**Publisher:** Fabien Savenay  
**Editor:** Paul Smaglik  
**Sales Director:** Ben Crowe

### European Head Office, London

The Macmillan Building  
4 Crinan Street  
London N1 9XW, UK  
Tel +44 (0) 20 7843 4961  
Fax +44 (0) 20 7843 4996  
e-mail: [naturejobs@nature.com](mailto:naturejobs@nature.com)

### Group European Manager:

Leonie Wells (4954)

### European Manager:

Nevin Bayoumi (4978)

### UK/ RoW/ Ireland:

Matt Powell (4953), Ben Corp (4974),

Andy Douglas (4975)

**Holland/ Italy:** Nevin Bayoumi (4978)

**Scandinavia:** Sille Opstrup (4994)

**Spain/ Portugal:** Leonie Wells

(4954)

### Production Manager: Billie Franklin

To send films and materials use

London address above.

Tel +44 (0) 20 7843 4814

Fax +44 (0) 20 7843 4996

e-mail: [naturejobs@nature.com](mailto:naturejobs@nature.com)

### International

#### Advertising Coordinator:

Laura Pearson (4977)

#### Naturejobs web development:

Tom Hancock

#### Naturejobs online production:

Ben Lund

### European Satellite Offices

#### France/ Belgium:

Christine Niox-Chateau

Tel + 33 (0) 1 43 20 16 51

Fax + 33 (0) 1 43 20 51 52

e-mail: [c.nioxchateau@nature.com](mailto:c.nioxchateau@nature.com)

#### Germany/ Austria/ Switzerland:

Patrick Phelan/ Kate Turner

Tel + 49 89 54 90 57 11/-2

Fax + 49 89 54 90 57 20

e-mails: [p.phelan@nature.com](mailto:p.phelan@nature.com)

[k.turner@nature.com](mailto:k.turner@nature.com)

### US Head Office, New York

345 Park Avenue South,

10th Floor, New York, NY 10010-1707

Tel +1 800 989 7718

Fax +1 800 989 7103

e-mail: [naturejobs@natureny.com](mailto:naturejobs@natureny.com)

#### US Sales Director: Ben Crowe

#### US Sales Manager: Peyton Mason

#### US Advertising Coordinator:

Comisa Salzman

### Japan Head Office, Tokyo

MG Ichigaya Building (5F),

19-1 Haraikakamachi,

Shinjuku-ku,

Tokyo 162-0841

Tel +81 3 3267 8751

Fax +81 3 3267 8746

e-mail: [k.cowan@naturejp.com](mailto:k.cowan@naturejp.com)

#### Japan Manager: Katie Cowan

# naturejobs

## Big or small neuroscience?

Over the past few years, a number of large multidisciplinary neuroscience centres have sprung up around the world. And there are several more on the drawing board, including what may well be a leading example for the field — a \$270-million facility on the Maryland campus of the US National Institutes of Health.

Large-scale facilities are important because the computational and imaging equipment needed in neuroscience is expensive and should be shared. In addition, bringing scientists from various disciplines together in one place and organizing them around broad research areas should help to advance the field by providing different perspectives on a given problem. So says one of the plan's architects, Gerry Fischbach, Columbia University's vice-president for health and biomedical sciences.

But Fischbach doesn't yet know what the plan, or others like it, will mean for neuroscience employment. Ideally, such centres would increase the overall opportunities for neuroscientists. But, if they are improperly funded, they might result in just the reshuffling of existing personnel.

Either way, Nancy Rothwell, president of the British Neuroscience Society and professor of biology at the University of Manchester, agrees that there is a move towards "larger critical mass" in neuroscience. But she cautions that such efforts shouldn't mean that scientists who prefer to go it alone are excluded from funding. "Many here would not want to see the loss of the individual scientist who is clearly outstanding and may not want or need to be part of a larger group," she says.

It's clear that scientific arguments exist for both approaches. But all too often, political and economic considerations prevail — disappointing proponents of either or both.

### Paul Smaglik

Naturejobs editor



## Contents

### CAREERS AND RECRUITMENT

The US adopts a multidisciplinary approach to neuroscience p4

European neuroscience — growing but not yet integrated p6

Long-term issues in Japan p8

### WWW.NATUREJOBS.COM

Career centre  
Information on the scientific job market

#### FOCUS

#### SPOTLIGHT

#### RECRUITMENT

#### SCIENTIFIC ANNOUNCEMENTS

#### SCIENTIFIC EVENTS



# All systems go for neuroscience

Multidisciplinary approaches pave the way towards new frontiers in understanding complex human behaviour and intractable diseases. Diane Gershon assesses the US field.

Exposed: the behaviour of nerve cells are coming under intense scrutiny.

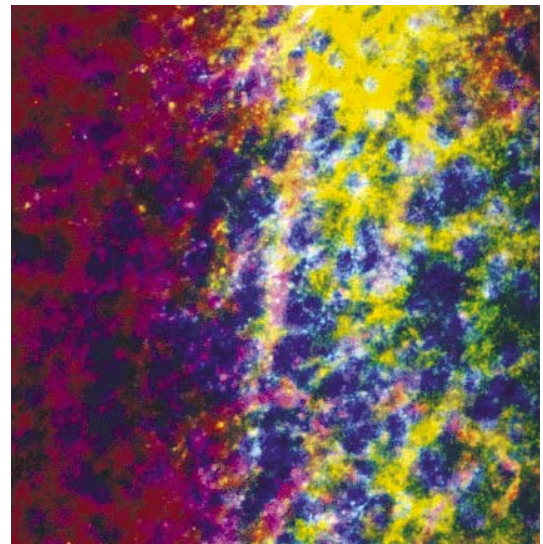
Interest in neuroscience has risen dramatically in the past 30 years. Membership of the US Society for Neuroscience has soared from 500 when it was founded in 1969, to more than 28,000 today. And fresh neuroscience initiatives — be they institutes (see 'Unlocking the secrets of autism', opposite), centres or looser associations of individuals — have been cropping up all around North America. Most new efforts reflect the growing feeling that better treatment for disorders of the nervous system and greater understanding of complex human behaviour will require a more integrated and interdisciplinary approach to neuroscience research and teaching.

## BUILDING BRIDGES

Carla Shatz, chair of the department of neurobiology at Harvard Medical School in Boston, sees the opportunities and challenges for neuroscience as having reached the 'integrative' neuroscience level. She regards the first 30 years of the field as having been largely reductionist. "I think the challenge now is how to put the molecules back into the cells, and the cells back into the systems, and the systems back into trying to really understand behaviour and perception," Shatz says. In order to accomplish this at Harvard, new connections are being built between the college and the medical school — two campuses that are physically separated by the Charles River.

A more integrative approach to neuroscience requires a technological infrastructure. And one important component of this is imaging, says Shatz. Accordingly, the medical school, in collaboration with Boston's Brigham and Women's Hospital, will open a new Brain Imaging Center this month dedicated solely to research. The centre will house a 4.7-tesla magnet for use in magnetic resonance imaging studies in mice and small non-human primates. It will also promote approaches, such as the use of confocal and two-photon microscopy, that allow researchers to look at molecules functioning *in vivo*.

The centre will serve as one of the core facilities for the new Harvard Center for Neurodegeneration and Repair. This large, distributed effort was jump-started recently after the medical school received a large anonymous gift, and so far it has pulled together more



C. SHATZ

than 100 clinicians and researchers.

The medical school is also considering whether to create a specific integrative physiology or neuroscience initiative — perhaps in the form of a new centre, says Shatz. The plans are not formalized and there is talk of broadening any such initiative beyond neuroscience, because the need to understand how systems function goes well beyond those of the brain, she says. In the meantime, the neurobiology department is expanding and will have four or five faculty slots to fill over the next few years.

On the Cambridge side of the river, plans are under way to establish a Center for Systems Neuroscience. "The goal of this initiative is to provide an umbrella for the systems neuroscience research that goes on and — ultimately — to provide a common laboratory building," says Markus Meister, a professor in the department of molecular and cellular biology.

Meister says that the formation of the centre does not mean there are plans to create a new department or to change existing departmental structures. But reallocation of space will allow individuals from biology, engineering and the applied sciences, physics and psychology to work side by side. A new building and vivarium are still four or five years away; when



Carla Shatz: facing the challenge of reassembling neuroscience.

completed, these will house 10–15 faculty. But Meister is not waiting for new premises to start recruiting. Some appointments have already been made at the junior faculty level, and there is an ongoing effort to attract a centre director.

### LOOKING NORTH

Neuroscience in general, and systems neuroscience in particular, has been a strength at Queen's University in Kingston, Ontario. But, until recently, the only formal recognition of this was a seminar series, says physiology professor Douglas Munoz. In January of this year, however, the university set up a Centre for Neuroscience Studies, with Munoz as its director, to promote integration of the neurological and psychiatric professions. Although in its early days, the initiative has already drawn together in a 'virtual' centre some 45 investigators dispersed about the campus and spanning both clinical and basic research.

Initially, the centre will focus on developing a new graduate-training programme in neuroscience, and establishing core facilities. It will also be used to drive strategic recruitments, particularly in the areas of computational neuroscience, molecular neurobiology, cognitive/behavioural neuroscience and clinical/rehabilitation neuroscience.

But, as it is a small university, Munoz acknowledges that Queen's will not become excellent in all areas of neuroscience and instead must focus on a few. Research clusters will be formed around 'research chairs' — the government programme aimed at recruiting and retaining top scientists to Canadian universities. Munoz, who was one of seven Queen's professors to be awarded a research chair in the first funding round, expects that 10 additional faculty will be recruited over the next couple of years.

Even before the new centre was formed, the sensorimotor systems group at Queen's, one of the larger neuroscience groups on campus, was already a cohesive group. Munoz, who belongs to this group, says that, as a PhD student, he was in a lab with one

principal investigator, one technician and one other grad student. "That's not at all the model I use to run my lab now," he says. Now, as one of seven principal investigators in the group, he has access to many technicians but pays only portions of their salaries.

Assembling large multidisciplinary research teams to work on a particular neurological problem is all in a day's work for William Luttge, director of the McKnight Brain Institute at the University of Florida in Gainesville, which opened new premises in 1998. On a fairly regular basis, Luttge must deal with 10 or so deans, 50 department chairs and some 325 faculty — a task that is made somewhat easier by having all the professional schools on one campus.

But, as a past chairman and dean himself, he is also aware of the importance of diplomacy. The McKnight institute, for example, does not dictate how graduate programmes should be run. And, even though an individual may receive all or part of their salary from the institute, they still report to a departmental chair. Issues of promotion and tenure are also handled at the departmental level. Luttge says that his role — and that of the institute — is instead to create opportunities, once an area of mutual research interest has been identified, and to remove barriers.

"I try very hard, when we are looking at programmes, to include as many people as possible, not in an egalitarian sense, but to embolden and empower the various departments and colleges," says Luttge.

Space — or lack of it — is the biggest problem now for Luttge. The institute has plans to build a \$100-million Neuro-Clinical Research Center adjacent to Shands at the University of Florida, the university's teaching hospital, but this is a longer-term aim, he says.

Diane Gershon is assistant editor of *Nature Medicine*, *New Technology*.

### Web links

MIND Institute ♦ [mindinstitute.ucdmc.ucdavis.edu](http://mindinstitute.ucdmc.ucdavis.edu)  
International meeting for Autism Research ♦ [www.imfar.org](http://www.imfar.org)  
McKnight Brain Institute ♦ [www.mbi.ufl.edu](http://www.mbi.ufl.edu)  
Center for Neuroscience Studies ♦ [www.queensu.ca/neurosci](http://www.queensu.ca/neurosci)

## Unlocking the secrets of autism



Research director of the MIND Institute, David Amaral (right), plans to get to the heart of neurodevelopmental disorders.

Pressure from parents who have children with autism has helped to boost the neuroscience profile at the University of California, Davis. After parents of autistic children in the Sacramento area demanded that more resources be devoted to understanding and treating the disorder, the Medical Investigation of Neurodevelopmental Disorders (MIND) Institute was launched in 1998. Besides autism, the institute studies other neurodevelopmental disorders, such as Asperger's syndrome, cerebral palsy and dyslexia.

Before the university formed a neuroscience centre about 10 years ago, "it had a pretty meagre community of neuroscientists," says David Amaral, research director of the MIND Institute. In the late 1980s, there were only about half a dozen neuroscientists, compared with 55 campus-wide today.

In September, the university broke ground on a new 135,000-square-foot building complex for the MIND Institute. The buildings are slated for completion by 2003 and will include wet and dry labs, as well as a clinic.

At full strength, the institute will have 20 scientists. Some preliminary appointments have already been made. But the real ramp-up in recruitment won't start until the new premises are nearer completion, says Amaral. The team will eventually consist of people with neuroimaging expertise, along with molecular neurobiologists. And, as there is a strong implication that children with autism have impaired immune systems — as some children have food allergies — Amaral also expects to recruit neuroimmunologists and nutritionists. **D.G.**





# A grassroots revolution

America may be thinking big, but in Europe the rise of neuroscience is finding favour in a more localized manner. Helen Gavaghan taps into the continental flavour of European interdisciplinary research.

**E**uropean scientists, much like their Japanese and US colleagues, see neuroscience as a frontier field in which multidisciplinary collaborations will lead to advances in understanding and therapeutics. But despite that shared interest, European countries generally approach the discipline from a different perspective, and on a different scale.

In Europe the interdisciplinary trend is mostly manifest when viewed from the grassroots level, rather than in the context of specific initiatives or new centres such as those in the United States (see 'All systems go for neuroscience', pages 4–5) or Japan (see 'Limited opportunity', page 8), says Monica di Luca, professor of pharmacology at the University of Milan and secretary-general of the Federation of European Neuroscience Societies (FENS).

FENS was established some six years ago to provide a single European forum for all the subjects involved in understanding the structure and function of the brain, and to promote neuroscience in Europe. It does not replace all the individual national societies, which tackle issues relevant to their countries.

The first FENS meeting attracted only about 2,000

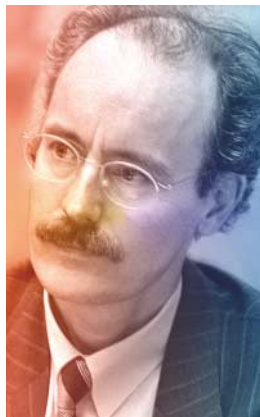
participants; the third, to be held next year in Paris, looks set to attract between 6,000 and 7,000. Although not as huge as the meetings of the US Society for Neuroscience, FENS is taking a grassroots approach to make itself felt as a focal point for European neuroscience.

That strategy appears to be working. "We are seeing a growing interest from young people in neuroscience," says di Luca. "We answer perhaps 800 e-mail queries per month about training, conferences and bursaries."

## DIVERSE RESULTS

FENS appears to be creating more general interest in neuroscience throughout Europe. But the grassroots approach makes it difficult to generalize further about European approaches to neuroscience or opportunities in the field, as situations are different in each of the countries across the continent.

For example, competition for postdoctoral positions remains fierce in England, whereas opportunities abound for postdocs wanting to work in Germany. And multi-site collaborations appear to be



Hans-Peter Thier: taking a multidisciplinary approach.

## Paths in neuroscience and advice

Nancy Rothwell and Peter Sneddon took different paths in their neuroscience careers, but they offer common advice.

After a series of fellowships, Rothwell became president of the British Neuroscience Association and a biology professor at the University of Manchester. Sneddon, on the other hand, decided after 16 years in academia to look for a different challenge. He found it at the Wellcome Trust, where, as head of the neuroscience division, he enjoys a broad overview of the field.

Rothwell and Sneddon both believe their perspective on finding a career in neuroscience applies to students and postdocs both inside and outside the United Kingdom.

They agree that the first hurdle is to choose the right lab for your PhD. Students, for example, need to be careful that they do not become just another part of the lab's labour force.

Then comes the search for a postdoctoral position. Rothwell's advice is to begin the search in good time. She doubts that the best students will be

looking for jobs in classified ads, and recommends getting known by networking at conferences.

Think about what it takes to climb the academic ladder, says Sneddon. He remembers that those of his contemporaries who went on to head labs were not content simply to be an extra pair of hands in the lab.

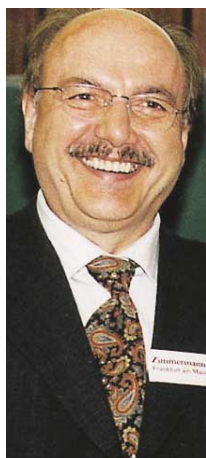
Rothwell recalls one candidate who came into the interview with a list of questions about every aspect of the lab, from the publication record to the supervision of graduate students. "I



Nancy Rothwell: has an international view.

employed that person immediately," she says.

H.G.



**Herbert Zimmermann:** laments that Germany finds it hard to attract neuroscientists.

more common in Germany than they are in France (see 'Breaking the boundaries', below).

Despite that diversity, some common threads remain. At a societal level, the pressures driving neuroscience in continental Europe are similar to those elsewhere: the sheer awfulness, for example, of neurodegenerative disease, and the likelihood that such diseases will become increasingly prevalent as the population ages.

"These are about the worst things that can happen to a human being," says Herbert Zimmermann, professor at the University of Frankfurt and

president of the German Neuroscience Society, which caters to basic rather than applied neuroscience.

And, like their Japanese and US counterparts, European scientists recognize that making inroads against that awfulness will require a quintessentially interdisciplinary approach. "In my lab," says Hans-Peter Thier, professor of cognitive neuroscience at the University of Tübingen, "I have clinicians, physicists, computer specialists and biologists."

New relationships are constantly being formed. "We're seeing a lot of applications from mouse genome people exploring links between genetics and behaviour," says Peter Sneddon, head of the neuroscience division at the Wellcome Trust.

New techniques are also creating new opportunities. One of the hottest areas in neuroscience, according to Zimmermann, is cognitive neuroscience, in which non-invasive imaging techniques are seen as a potential way to understand the link between the brain's functioning and human behaviour.

#### GERMAN IDEAL

Within Europe, Germany has perhaps the biggest head start in integrating different aspects of, and approaches to, neuroscience. Its universities have strong internationally recognized groups, with plans for more courses and groups on the way.

Germany also comes closest to the model in the United States and Japan of having dedicated institutes rather than disparate labs. The Leibniz Institute in Magdeburg and an expanding Max Planck Institute at Frankfurt are both completely devoted to neuroscience.

Yet despite this vibrancy in German neuroscience, the country finds it difficult to attract good postdocs. Zimmermann considers that the German system has much to do with the difficulty.

Until recently, there were no tenure-track positions, and even now there are only a few. Institutions do not allow postdocs to stay at an institution for more than five years, because those who stay longer become entitled to full-time benefits. Even postdocs whose work is admired, and who look likely to achieve a full-time post, are encouraged to go elsewhere rather than stay at their institution.

Although postdocs elsewhere have their own

difficulties, this absolute limit of five years and lack of tenure-track posts does discourage young people from following a scientific career. The situation is not unique to neuroscience, but Zimmermann believes it is more acute in this field than in, say, molecular biology, where people have a wider choice of career in industry once their five years are up.

Despite the problems, Germany is pressing ahead with plans to strengthen its neuroscience activities. For example, the DFG — the country's central research-funding body — is currently evaluating proposals for a centre of excellence in neuroscience that will be established at an existing university. The call for proposals was entitled "From molecules to cognition". The winner will be announced next May.

"The winner will receive DM10 million (\$4.6 million) per year to fund new posts for full professors, assistant professors and technical support staff," says Sabine Behrenbeck, a programme director at the DFG.

Helen Gavaghan is a freelance writer based in West Yorkshire.

#### Web links

Federation of European Neuroscience Societies ♦ [www.fens.org](http://www.fens.org)

Wellcome Trust ♦ [www.wellcome.ac.uk](http://www.wellcome.ac.uk)

The European Dana Alliance for the Brain ♦ [www.edab.net](http://www.edab.net)

## Breaking the boundaries

**Martin Giurfa is in a good position to compare trends in French and German neuroscience. The professor of behavioural neurobiology at Paul Sabatier University in Toulouse spent 11 years at the Free University of Berlin's Institute of Neurobiology before moving to France this year.**

**He says that, in general, French scientists tend to hire from within the country. He's hoping to make his Laboratory of Animal Cognition more international, by maintaining the collaborations he started in Germany and the United States, as well as creating new ones within France and other countries.**

**Increasing the laboratory's international profile will in turn help attract good neuroscientists from outside the country, he**

**says. But it will be necessary to overcome some cultural and scientific quirks.**

**For example, in Germany, some universities conduct advanced courses in English, even if there are only Germans in the classrooms. This strategy prepares the students for the international market and makes the courses more attractive to foreigners. "That would never happen in France," he says.**

**And government funding tends to keep vertebrate and invertebrate animal models separate. He'd like to see that division — as well as the cultural tendency towards isolation and away from English — end. He suspects that as the European Union plays a bigger role in science funding, some of these boundaries will dissolve.**

Paul Smaglik



# Limited opportunity

Japan's Brain Science Institute offers young neuroscientists jobs — but doesn't guarantee them long-term employment, says Robert Triendl.

**T**he Brain Science Institute (BSI) at RIKEN, Japan's Institute of Physical and Chemical Research, is a magnet for young Japanese neuroscientists. But some young neuroscientists worry that the emphasis on limited-term contracts at the BSI does not provide them with long-term stability.

Founded in 1997, the BSI is the first Japanese neuroscience research institute dedicated exclusively to basic science. It employs 600 researchers in 35 research laboratories that span all fields of neuroscience. Located at RIKEN's Wako campus, some 30 minutes from central Tokyo, the BSI offers working conditions and funding opportunities that far exceed what scientists can expect in Japan's academic environment.

But Japan's finance ministry has kept tight control on the creation of any new faculty positions at Japan's national research institutes or universities — which means that all scientists at the BSI have a five-year contract that, in principle, is not renewable. Scientists who have not been promoted to the position of group leader during the five years have to leave the institute.

Limited-term positions are especially problematic for female scientists. Although more female researchers are joining the BSI, there are virtually no female group leaders. Women scientists in Japan have generally had a difficult time rising to top positions (see *Nature* **410**, 404–406; 2001). Nevertheless, some women at the BSI say they prefer the short-term contract there compared with the difficulty they face getting a permanent university position.

Masao Ito, director of the institute, which he helped to establish, defends the limited-term policy, arguing that it is

meant “to increase the mobility of young researchers”. But as the pool of established neuroscience researchers in Japan is fairly small, the BSI faces increasing competition for senior faculty from large national universities.

Earlier this year, an international advisory panel urged the BSI to consider setting up a number of long-term faculty appointments to ensure that the institute keeps the most successful scientists. As universities gain more independence from central government over the next few years, competition for successful faculty will only increase.

The BSI's attempts to negotiate a more equal distribution of tenured positions with RIKEN's various institute laboratories have so far yielded little success. All of the 650 tenured positions at RIKEN are held by scientists in traditional institutes, rather than newer set-ups such as the BSI.

The BSI's emphasis on limited-term contracts is intertwined with its history, which goes back to the ‘frontier research system’ — a special research-funding scheme set-up by the government and RIKEN in the mid-1980s. The well-funded programme was meant to attract foreign scientists for temporary periods. Ito took over the directorship of that programme, and used it as a basis to expand neuroscience research and to launch the BSI. Many of the neuroscience laboratories established within the frontier programme later moved into the adjacent new building of the BSI located at the same campus. But finding a way to overcome that history, by allowing some tenured positions within the BSI, may well give the institute — or at least the scientists working in it — a brighter future. ■

Robert Triendl is a freelance writer based in Tokyo.

## Web links

Brain Science Institute  
 ♦ [www.brain.riken.go.jp](http://www.brain.riken.go.jp)  
 National Institute of Neuroscience  
 ♦ [www.ncnp.go.jp/nin](http://www.ncnp.go.jp/nin)  
 Mitsubishi Institute of Life Sciences  
 ♦ [www.m-kagaku.co.jp/english/aboutmcc/RD/relation/life.htm](http://www.m-kagaku.co.jp/english/aboutmcc/RD/relation/life.htm)  
 Neuroscience Research Institute  
 ♦ [unit.aist.go.jp/neurosci/english/english.htm](http://unit.aist.go.jp/neurosci/english/english.htm)  
 Tokyo Metropolitan Institute for Neuroscience  
 ♦ [www1.sphere.ne.jp/tmin](http://www1.sphere.ne.jp/tmin)

## A community takes root

By US standards, Japan's neuroscience community is still in its infancy. The Japanese Neuroscience Society has only a few thousand members and, despite much recent growth, its annual meetings have preserved a certain intimacy.

The first dedicated research centres, such as the National Institute of Neuroscience in Tokyo, were conceived mostly as research services to hospitals or medical facilities. Other

institutions, such as the Mitsubishi Kasei Institute of Life Sciences, the first large-scale life-sciences research centre in Japan's private sector, have also provided institutional niches for research in molecular neuroscience. Set up in 1971 by Mitsubishi Chemical, the institute was conceived as a long-term contribution to the development of life sciences in Japan, rather than as a corporate research unit.

R.T.